

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
16 September 2004 (16.09.2004)

PCT

(10) International Publication Number
WO 2004/078983 A2

- (51) International Patent Classification⁷: C12N 15/82, 15/29, 15/63
- (21) International Application Number: PCT/EP2004/002096 ✓
- (22) International Filing Date: 3 March 2004 (03.03.2004) ✓
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 03005181.7 ✓ 7 March 2003 (07.03.2003) ✓ EP
- (71) Applicant (for all designated States except US): BASF PLANT SCIENCE GMBH [DE/DE]; 67056 Ludwigshafen (DE).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): HOFVANDER, Per [SE/SE]; Doppinggränd 8, S-239 41 Falsterbo (SE). ANDERSSON, Mariette [SE/SE]; Pilegatan 20, S-222 29 Lund (SE). ✓
- (74) Agent: PRESSLER, Uwe; BASF Aktiengesellschaft, 67056 Ludwigshafen (DE).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published: — without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ENHANCED AMYLOSE PRODUCTION IN PLANTS

(57) Abstract: The invention relates to methods for increasing the amylose content in plants, preferably in potato plants, by expressing a starch biosynthesis enhancing protein. The invention furthermore relates to an expression cassette expressing the polypeptide in potato plants, preferably in the tubers, the transgenic plants expressing the polypeptide and to the use of said transgenic plants for the production of fine chemicals especially other than native starches.

WO 2004/078983 A2

54331
021017

k

THIS PAGE BLANK (USPTO)

Enhanced amylose production in plants

10/547 660

JC05 Rec'd PCT/PTO 01 SEP 2009

Description

- 5 The present invention relates to new starch biosynthesis enhancing proteins, nucleic acids encoding a starch biosynthesis enhancing protein, a method for producing amylose with high efficiency by culturing genetically modified plants with an increased amylose biosynthesis compared to the wild type or to the genetically modified plants themselves as well as the use of these transgenic plants over-expressing at least one
10 of the starch biosynthesis enhancing proteins for the production of amylose.

15 Starch is the major storage carbohydrate of plants and is mainly accumulated in seeds and tubers, which are then the reproductive tissues of plants that form those types of organs. Starch is also accumulated on a diurnal basis where starch is built up in green tissue from photosynthetic products and then metabolised for energy during the dark period. The storage starch is assembled into semi crystalline granules. Amylopectin and amylose are the two constituent molecules of starch. Amylopectin is a branched molecule consisting of linear α -1,4 glucan chains linked by α -1,6 bonds. Amylose consists essentially of the linear α -1,4 glucan chains.

20 Starch is utilised for many applications within the technical industry as well as the food industry. Main crops used by starch processors are maize and potato. For potato specific varieties are utilised for starch production that have been bred for high starch contents. This means that the starch content and yield is an important economic driver
25 for the starch processing industry. A greater part of produced dry starch is used for paper production. The specifications and requirements for the starch component varies from application to application and starch is many times chemically modified in order to provide desired properties to an application. Another way to achieve starch of different qualities is to take advantage of mutations in the starch biosynthesis and more recently
30 by genetic modification of pathways leading to starch. The first main modifications have been to separate the production of the two starch components amylopectin and amylose into different varieties. Waxy or "amylose free" varieties contain solely amylopectin type starch while there are also high amylose genotypes such as "amylose extender" in maize.

35 Amylose starch has several potential industrial uses as a film former or for expanded products. High amylose starch can be achieved in potatoes and other starch containing plants by inhibition of starch branching enzymes. This leads then to the concomitant reduction or elimination of amylopectin branching and thereby an increased amylose
40 fraction.

US 5,856,467 describes the genetically engineered modification of potato for suppressing formation of amylopectin-type starch. The document describes an antisense construct for inhibiting, to a varying extent, the expression of the gene coding for formation of starch branching enzyme (SBE gene) in potato, said antisense construct comprising a tuber specific promoter, transcription start and the first exon of the SBE gene, inserted in the antisense direction.

US 6,169,226 relates to an amino acid sequence of a second starch branching enzyme (SBE II) of potato and a fragment thereof as well as to the corresponding isolated DNA sequences. It describes the production of transgenic potatoes and the use of these transgenic potatoes for the production of amylose-type starch.

WO 97/20040 and WO 98/20145 describe methods of altering the amylopectine/amylose starch content of plant cells by introducing into the plant cells nucleic acid sequences operably linked in sense or antisense orientation to a suitable promoter which homologous genes encodes polypeptides having SBE I or SBE II activity.

A side effect of the amylose overproduction is a decreased total starch content in the potatoes. This decrease becomes more pronounced as the amylose fraction is increased.

Basic enzymes for the production of amylopectin and amylose are starch synthases that build the linear α -1,4 glucan chains and branching enzymes breaking the α -1,4 glucan chain and reattaching them by α -1,6 bonds. Several other enzymes are likely to affect starch structure and composition, such as debranching enzymes, but initially most focus has been towards affecting the expression of starch synthases and starch branching enzymes. This has led to an extensive dissection of what enzymes are important for what features of starch synthesis. However it has never been convincingly shown how the synthesis of starch in plants whether amylose or amylopectin is initiated.

Suggestions on the initiation of starch biosynthesis have been the subject of several scientific papers since it has been difficult to attribute a primer independent function to starch synthases under other than artificial in vitro conditions. By primer independent function implies the formation of new α -1,4 glucan chains with ADP-glucose as the sole starting point and building block. One proposed pathway has been that the presence of maltooligosaccharides act as primers for the addition of further glucose units by starch synthases although it has been debated on whether concentrations are sufficient to provide the basis for starch synthesis and also how these maltooligosaccharides would be formed in the plastids.

Starch is in plants synthesised as an energy storage molecule. Much is known about the enzymes participating in the starch biosynthesis although, the initiation of the starch molecule has remained unsolved. In mammals and yeast an energy storage molecule very similar to starch is synthesised, glycogen. The enzymatic steps for synthesis of the respective molecules are analogous. In glycogen biosynthesis the initiation of the molecule is known and synthesised by the enzyme glycogenin. Glycogenin is a self-glucosylating enzyme polymerising a linear chain of approximately 8 glucose molecules on itself. The primer of about 8 glucose residues is necessary for the enzymes catalysing the continuation of glucose incorporation to the glycogen molecule to function.

Cheng et al., 1995, Mol. and Cell. Biol. 6632-6640 compare the two yeast proteins with rabbit muscle glycogenin.

Roach et al., 1997, Progress in Nucleic Acid Research and Molecular Biology Vol 57, describe self glycosylating initiator proteins and their roll in glycogen biosynthesis.

Mu et al., 1997, Journal of Biological Chemistry 272 (44), 27589-27597 compare mammalian with yeast and *C. elegans* glycogenins.

Factors important for starch quantity have been investigated and many initiatives have been taken, especially in potato, to increase starch formation and content by over-expression or inhibition of various enzyme activities in areas of increased substrate supply, increased biosynthesis activity or shutting down substrate diverting pathways but so far this has led only to limited success with no commercial applications and only some scientific publications.

Regierer, B. et al., Starch content and yield increase as a result of altering adenylate pools in transgenic plants. Nat Biotechnol. 20(12):1256-60, (2002).

Sweetlove, LJ et al., Starch synthesis in transgenic potato tubers with increased 3-phosphoglyceric acid content as a consequence of increased 6-phosphofructokinase activity. Planta 213(3):478-82 (2001).

Veramendi, J et al., Antisense repression of hexokinase 1 leads to an overaccumulation of starch in leaves of transgenic potato plants but not to significant changes in tuber carbohydrate metabolism. Plant Physiol. 121(1):123-34 (1999).

Geigenberger, P et al., Overexpression of pyrophosphatase leads to increased sucrose degradation and starch synthesis, increased activities of enzymes for sucrose-starch

interconversions, and increased levels of nucleotides in growing potato tubers. *Planta*. 205(3):428-37(1998).

- 5 Sweetlove, LJ et al., Starch metabolism in tubers of transgenic potato (*Solanum tuberosum*) with increased ADPglucose pyrophosphorylase. *Biochem J.* 320 (2):493-8 (1996).

10 In other research a biochemical function superficially similar to the one initiating glycogen production in animals was investigated. A class of genes have then been isolated from several plants and was given the name amylogenin (WO94/04693; Sing, D. et al, β -Glucosylarginin: a new glucose-protein bond in a self-glucosylating protein from sweet corn, *FEBS Letters* 376:61-64, (1995) in the belief that it was the plant equivalent of glycogenin which acts as a self-glucosylating enzyme and provide primers for starch biosynthesis in plants. These genes have no resemblance from a structural point
15 of view to the genes coding for glycogenin and have later been determined not to have a function in starch biosynthesis but rather might be of importance for cell wall formation, see Bocca, S.N et al., Molecular cloning and characterization of the enzyme UDP-glucose: protein transglucosylase from potato. *Plant Physiology and Biochemistry* 37(11):809-819(1999).

20 WO 98/50553 describes nucleic acid fragments encoding a plant glycogenin or a water stress protein. WO 98/50553 also relates to the construction of chimeric genes encoding all or a portion of a plant glycogenin in sense or antisense orientation, wherein expression of the chimeric gene results in production of altered levels of a plant glycogenin in a transformed host cell.
25

Thus although many enzymes and pathways have been investigated in plants, the question on how starch formation is initiated and what determines the starch content is still unresolved.
30

Amylose is a commercially important starch product with many uses but unfortunately an increase in amylose content in transgenic potato plants is associated with a significant decrease in starch content, see figure 1.

35 Analyses of transgenic high amylose potato lines show that there is an excess of soluble sugars in these lines, see figure 2. This indicates that the starch biosynthesis in these transgenic lines is not efficient enough for incorporation of available sugars.

40 Amylose starch consists of very few reducing ends compared to native starch. Therefore it is commercially important to identify genes that further enhance the amylose

biosynthesis and that are capable to incorporate the excess of glucose residues available and to compensate the decrease in starch content in plants that produces amylose in high amounts.

- 5 The invention aims at enhancing the yield of amylose biosynthesis by the over-expression of genes which enhance starch biosynthesis in transformed plants.

The invention describes genes coding for proteins which enhance starch production.

- 10 The present invention describes the nucleic acids SEQ ID NO 1 and 3 from potato coding for enzymes enhancing the de novo starch biosynthesis.

- Example 1 describes that the nucleic acid sequences SEQ ID NO 1 or 3 can complement a missing glycogenin function in yeast cells containing knock-out mutations for the self-glycosylating proteins Glg1p and Glg2p.
- 15

- Gene constructs were made for gene-inhibition and over-expression of the two genes SEQ ID NO 1 or 3 in potato. Transgenic lines with the over-expressed or inhibited enzyme activity were analysed with regard to the genes influence on starch content.
- 20

Both genes were inserted in sense and antisense direction downstream of a plant promoter element, resulting in the transformation binary vectors pHS1, pHS2, pHS3 and pHS4, see figures 3-7.

- 25 The antisense constructs were transformed into the potato plant varieties Prevalent and Producent and the sense constructs were transformed to the potato variety Desiree and the transgenic plant AM99-2003 according to the transformation method as described in example 2. The transgenic plant AM99-2003 was produced as described in example 3.

- 30 Prevalent and Producent are starch varieties having a starch content of approximately 20 %. Desirée is a potato variety having a starch content of approximately 16% and AM99-2003 is a transgenic high amylose line having a starch and thereby amylose content of approximately 13%.

- 35 The putative genes were isolated from a tuber specific cDNA library of *Solanum tuberosum* (variety Prevalent). The library was made from a lambdaZAP directional kit (Stratagene).

Both cDNAs isolated were full-length clones of the individual genes and named StGH1 and StGH2, for nucleic acid sequences see SEQ ID NO 1 and SEQ ID NO 3.

pHS1

5

A 1300bp PCR fragment from the StGH1 gene was constructed in antisense direction driven by the gbss promoter. The PCR fragment was cut out from its cloning vector pCR4-TOPO (Invitrogen) with EcoRI (blunted) and XbaI. The fragment was ligated to the pGPTV-kan (Becker, D. et al., Plant Molecular Biology 20:1195-1197(1992)) based binary vector pHo3.1 between a gbss promoter (WO 92/11376) and a nos terminator at the Sall (blunted) and XbaI sites. The binary vector also includes nptII as selection marker driven by the nos promoter (Herrera, L. et al., 1983). The construct was named pHS1, for details see figure 3a and 4.

15 pHS2

A 2300bp full-length cDNA clone of StGH2 was cut out from the cloning vector pBluescript (Stratagene) with XbaI and XhoI. The gene was ligated in antisense direction between the gbss promoter and nos terminator to the binary vector pHo3.1 at XbaI and Sall. As can be seen under pHS1 the vector has nptII as selection system. The vector was named pHS2, for details see figure 3b and 5.

pHS3

25 A full-length StGH1 cDNA, (1780bp) was cut out from the host vector pBluescript with EcoRI (blunted) and BglII and ligated to the BamHI and SmaI sites of pUCgbssprom (3886bp), containing pUC19 with the gbss promoter and the nos terminator. The plasmid was named pUCGH1.

30 A fragment with the gbss promoter, the StGH1 gene and the nos terminator was moved from pUCGH1 with EcoRI (blunted) and HindIII (2980bp) and ligated to PstI (blunted) and HindIII opened pSUN1 (WO 02/00900). The plasmid was named pSUNGH1.

35 A 3600bp fragment containing the AHAS resistance gene from Arabidopsis thaliana (Sathasivan, K. et al., Plant Physiology 97(1991), 1044-1050) with nos promoter, see Herrera-Estrella, L. et al., Nature 303:209-213(1983) and OCS terminator (Wesley, S.V. et al., Plant J. 27(6):581-590(2001)) was ligated to pSUNGH1 (9000bp) at the SmaI site. The vector was given the name pHS3, for details see figure 3c and 6.

pHS4

The gbss promoter and nos terminator was ligated to pBR322 with EcoRI and HindIII. Between the promoter and terminator an EcoRI-HincII full-length gene pStGH2 was
5 cloned at the XbaI site. The 3366bp promoter-gene-terminator complex was cut using EcoRI (partial digestion) and EcoRV, and ligated to pSUN1 at EcoRI-EcoRV and named pSUNGH2. An XbaI fragment with AHAS gene (*Arabidopsis thaliana*), nos promoter and OCS terminator was ligated to pSUNGH2 opened with XbaI (partial digestion). The AHAS gene is used as selection marker. The construct was named
10 pHS4, for details see figure 3d and 7.

Example 2 describes the general method for the transformation of different potato plant varieties producing native starch or high amylose type starch with pHS1, pHS2, pHS3 or pHS4.

15 The StGH1 and StGH2 genes were down-regulated in the potato plant varieties Prevalent and Producent by transformation with the genes in antisense direction in relation to a plant regulatory element as described in example 4 and 6. Down-regulation of the two genes resulted in a decrease in gene expression in transgenic lines compared to
20 their mother varieties in the order of 50-95%, see example 7 and table 3. Transgenic lines transformed with pHS1 and pHS2 with confirmed decrease in gene-expression have a decrease in dry matter of 7 to 11% compared to their mother varieties, see example 8 and table 5.

25 The StGH1 and StGH2 genes were over-expressed in potato driven by the tuber specific promoter gbss, as described in example 5. A mutated AHAS gene was used as selection marker yielding tolerance to the Imazamox herbicides. Two potato varieties were transformed, Desiree and AM99-2003 a transgenic high amylose line with a 40% decrease in starch content compared to its parental line. The transformed lines over-
30 expressing StGH1 and StGH2 were selected as described in example 6. The gene expression levels were analysed with real-time PCR, see example 7 and table 3. The over-expression of the genes StGH1 and StGH2 resulted in a 2 to 10 times increase in gene expression compared to their parental line. Furthermore the lines over-expressing StGH1 and StGH2 showed an increase in dry matter of up to 36 % as described in
35 example 8 and table 5.

The over-expression of StGH1 and StGH2 in transgenic potato plants producing amylose type starch resulted in an increased dry matter content, which means an increased amylose content as no amylopectin is produced, see examples 8 to 12.

40

RNA interference (RNAi) functions by introduction of double stranded RNA (dsRNA) into a cell, which causes a degradation of the homologous RNA. The dsRNA is cleaved into small interfering RNA (siRNA) of 21-25 nucleotides by a ribonuclease called Dicer. The siRNA connects with a protein complex and forms a RNA-induced silencing complex RISC. The RISC becomes activated by ATP generated unwinding of the siRNA, which binds to the homologous transcript and cleaves the mRNA resulting in gene silencing, see Mc Manus MT and Sharp PA., Gene silencing in mammals by small interfering RNAs. Nature Rev Genet 3:737-747(2002);

- 10 Dillin A., The specifics of small interfering RNA specificity. Proc Natl Acad Sci USA 100(11):6289-6291 (2003);

Tuschl T., Expanding small RNA interference. Nature Biotechnol 20:446-448 (2002)

- 15 Production of high amylose lines was more efficient when using the RNAi constructs pHAS3 (figure 23) and pHAS8b (figure 20) than the antisense construct pHAbE12A. The frequency of high amylose lines of total transgenic shoots produced when using for example pHAS8b and pHAS3 is above >25%, compared to a frequency of approximately 1% high amylose lines of total transgenic shoots produced , see example 15 to 17.

- 20 The RNAi constructs pHAS8b (figure 20) and pHAS3 (figure 23) (SEQ ID NO 24) used for high amylose potato production contain a be1 and be2 fragment (SEQ ID NO 19) cloned in inverted tandem. The constructs are only differing in the spacer used located between the inverted repeats where for pHAS8b a fragment of the be2 promoter was used (SEQ ID NO 18) while for pHAS3 a cloning residue from pBluescript was used (SEQ ID NO 23). The RNAi constructs resulted in efficient down-regulation of the branching enzyme genes.

- 30 Furthermore, the fragments of respective be1 and be2 genes could be shorter or longer and could be targeting other parts of the branching enzyme genes. Shorter fragments for RNAi of be1 and be2 are described in SEQ ID NO 21 and 22.

- 35 The starch biosynthesis enhancing protein according to the invention comprises the amino acid sequence SEQ ID NO 2 or 4 or a protein which comprises a sequence derived from SEQ ID NO 2 or 4, which is at least 50%, preferably at least 60%, more preferably at least 70%, more preferably at least 80%, still more preferably at least 90%, most preferably at least 95%, identical at the amino acid level to the sequence

SEQ ID NO 2 or 4 and has the property of a starch biosynthesis enhancing protein.

This starch biosynthesis enhancing protein may also be prepared by artificial variations starting from the SEQ ID NO 2 or 4, for example by substitution, insertion or deletion of amino acids.

5

Such a protein can be used to increase the production of amylose or amylopectin in non-transgenic or transgenic plants.

10

The term "substitution" in the specification means the replacement of one or more amino acids by one or more amino acids. Preference is given to carrying out "conservative" replacements in which the amino acids replaced has a property similar to that of the original amino acid, for example replacement of Glu by Asp, Gln by Asn, Val by Ile, Leu by Ile, Ser by Thr.

15

"Deletion" is the replacement of an amino acid or amino acids by a direct bond. Preferred positions for deletions are the polypeptide termini and the junctions between the individual protein domains.

20

"Insertions" are insertions of amino acids into the polypeptide chain, with a direct bond formally being replaced by one or more amino acids.

25

"Identity" between two proteins means the identity of the amino acids over the in each case entire length of the protein, in particular the identity which is calculated by comparison with the aid of the Vector NTI Suite 7.1 Software of the company Informax (USA) using the Clustal W method (Thompson, JD et al., Nucleic Acid Research, 22 (22):4673-4680, 1994)

with the parameters set as follows:

30

Multiple alignment parameter:

35

Gap opening penalty	15
Gap extension penalty	6.66
Gap separation penalty range	8
Gap separation penalty	on
% identity for alignment delay	40
Residue specific gaps	on
Hydrophilic residue gap	off
Transition weighing	0

40

Pairwise alignment parameter:

FAST algorithm	off
K-tuple size	2
Gap penalty	5
Window size	4
Number of best diagonals	4

10 Accordingly, a protein which is at least 50% identical at the amino acid level to the sequence SEQ ID NO 2 or 4 means a protein which, when comparing its sequence with the sequence SEQ ID NO 2 or 4, is at least 50% identical, in particular according to the above program algorithm using the above set of parameters.

15 Further natural examples of genes coding for a starch biosynthesis enhancing protein according to the invention can readily be found, for example, in various organisms, in particular in plants, whose genomic sequence is known by comparing the identity of the amino acid sequences or of the corresponding back-translated nucleic acid sequences from databases with the sequence of SEQ ID NO 2 or 4, in particular according to the
20 above program algorithm using the above set of parameters.

In the completed genome sequence of *Arabidopsis thaliana*, five putative coding sequences can be deduced by searching for exon/intron boundaries and comparing with back translated sequences of SEQ ID NO 2 or 4.

25 The following nucleic acid sequences of *Arabidopsis thaliana* SEQ ID NO 5, SEQ ID NO 7, SEQ ID NO 9, SEQ ID NO 11 and SEQ ID NO 13 could be used to carry out the invention and are coding for the starch biosynthesis enhancing proteins SEQ ID NO 6, SEQ ID NO 8, SEQ ID NO 10, SEQ ID NO 12 and SEQ ID NO 14.

30 Furthermore the following nucleic acid sequences or ESTs can be used in order to identify and clone genes coding for a starch biosynthesis enhancing protein from plant organisms:

35 Tomato ESTs from GenBank: AW216407, BE450055, BF097262, BE450557, BF097173

Wheat ESTs from GenBank: BJ292476, BJ278875, BJ283925, BE442966, CA666180, BQ483228

40

Maize EST from GenBank: BG319971

Rice ESTs from GenBank: AL606633, CA752890, BI813265

- 5 Natural examples of starch biosynthesis enhancing proteins and the corresponding genes can furthermore readily be found in various organisms, in particular plants, whose genomic sequence is unknown by hybridization techniques in a manner known per se, for example starting from the nucleic acid sequences SEQ ID NO 1 or SEQ ID NO 3 or any of the SEQ ID NO 5, 7, 9, 11 or 13 or any of the EST sequences de-
10 scribed above.

The hybridization may be carried out under moderate (low stringency) or, preferably, under stringent (high stringency) conditions.

- 15 Such hybridization conditions are described, inter alia, in Sambrook, J., Fritsch, E.F., Maniatis, T., in: Molecular Cloning (A Laboratory Manual), 2nd edition, Cold Spring Harbor Laboratory Press, 1989, pages 9.31-9.57 or in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6.
- 20 By way of example, the conditions during the washing step may be selected from the range of conditions which is limited by those with low stringency (with 2X SSC at 50°C) and those with high stringency (with 0.2X SSC at 50°C, preferably at 65°C) (20X SSC: 0.3 M sodium citrate, 3 M sodium chloride, pH 7.0).
- 25 In addition, the temperature may be raised during the washing step from moderate conditions at room temperature, 22°C, to stringent conditions at 65°C.

- Both parameters, salt concentration and temperature, may be varied simultaneously and it is also possible to keep one of the two parameters constant and to vary only the
30 other one. It is also possible to use denaturing agents such as, for example, formamide or SDS during hybridization. In the presence of 50% formamide, the hybridization is preferably carried out at 42°C.

Some exemplary conditions for hybridization and washing step are listed below:

- 35 (1) hybridization conditions with, for example
- (i) 4X SSC at 65°C, or
- 40 (ii) 6X SSC at 45°C, or

- (iii) 6X SSC at 68°C, 100 mg/ml denatured fish sperm DNA, or
- 5 (iv) 6X SSC, 0.5% SDS, 100 mg/ml denatured fragmented salmon sperm DNA at 68°C, or
- (v) 6X SSC, 0.5% SDS, 100 mg/ml denatured fragmented salmon sperm DNA, 50% formamide at 42°C, or
- 10 (vi) 50% formamide, 4X SSC at 42°C, or
- (vii) 50% (vol/vol) formamide, 0.1% bovine serum albumin, 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 50 mM sodium phosphate buffer pH 6.5, 750 mM NaCl, 75 mM sodium citrate at 42°C, or
- 15 (viii) 2X or 4X SSC at 50°C (moderate conditions), or
- (ix) 30 to 40% formamide, 2X or 4X SSC at 42°C (moderate conditions).
- 20 (2) Washing steps of 10 minutes each with, for example
- (i) 0.015 M NaCl/0.0015 M sodium citrate/0.1% SDS at 50°C, or
- (ii) 0.1X SSC at 65°C, or
- 25 (iii) 0.1X SSC, 0.5% SDS at 68°C, or
- (iv) 0.1X SSC, 0.5% SDS, 50% formamide at 42°C, or
- 30 (v) 0.2X SSC, 0.1% SDS at 42°C, or
- (vi) 2X SSC at 65°C (moderate conditions).

35 Preferred proteins with starch biosynthesis enhancing activity are proteins from plants, cyanobacteria, mosses or algae, particular preferred from plants. A particular preferred protein comprises the amino acid sequence SEQ ID NO 2 or 4.

40 If, for example, the protein is to be expressed in a plant, it is frequently advantageous to use the codon usage of said plant for backtranslation and resynthesis of the gene according to codon usage of said plant.

The invention further relates to nucleic acids encoding a starch biosynthesis enhancing protein according to the invention. All of the nucleic acids mentioned in the specification may be, for example, a RNA sequence, DNA sequence or cDNA sequence.

5

Suitable nucleic acid sequences can be obtained, for example, by back-translating the polypeptide sequence according to the genetic code. For this, preference is given to using those codons which are used frequently according to the organism-specific codon usage. The codon usage can be readily determined on the basis of computer analyses of other known genes of the organisms in question.

10

All of the above-mentioned genes coding for a starch biosynthesis enhancing protein can furthermore be prepared in a manner known per se from the nucleotide building blocks by chemical synthesis, for example by fragment condensation of individual overlapping complementary nucleic acid building blocks of the double helix. The chemical synthesis of oligonucleotides may be carried out, for example, in a known manner according to the phosphoramidite method (Voet, Voet, 2nd edition, Wiley Press New York, pages 896-897). The annealing of synthetic oligonucleotides and filling-in of gaps with the aid of the Klenow fragment of DNA polymerase and ligation reactions and also general cloning methods are described in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press.

15

20

25

30

Genes coding for this function may be integrated in the plant chromosomes and upon expression utilize a transit peptide to localise to plastids which is the organelle where starch/amylose biosynthesis takes place or be integrated directly into the plastid genome and thereby surpass the need for the localisation signal. The genes may be expressed constitutively or organ specific. For organ specific expression, promoters with tuber specific expression is preferable in potatoes while in cereals as maize or wheat a endosperm specific expression would be preferred to achieve a high degree of expression in organs where storage starch is accumulated. When transformed to the plastid genome then specific regulatory elements suitable for that organelle apply.

35

The genes of this invention may be used in combination with other genes that can be situated on the same gene construct or transferred and combined by co-transformation or super transformation. Genes and traits that are of interest to combine with the genes of the instant invention are agronomic or input trait such as herbicide tolerance, disease and pest resistance or stress tolerance but could also be output traits such as starch structure modification or yield. Genes and traits used in combination with the genes described in the invention could be for adding a function that is not present in

the modified plant species or over-expressing a function that is already present or inhibiting a function by the use of antisense, RNAi or antibodies.

5 The invention may be used to increase the starch or amylose content in potato tubers but would in its context not be limited to potatoes but would be applicable to other starch producing and storing plants such as e.g. corn, cassava, wheat, barley, oat and rice.

10 The described invention is particularly suited for eliminating a lower starch content associated with increased amylose content in different plants where the number of α -1,4-glucan chain non-reducing ends is greatly reduced due to the reduction or elimination of α -1,6 branch formation. Amylopectin is an extremely efficient structure, as is glycogen, for polysaccharide production since it is very branched and thus contains as many points accessible for starch synthesis as there are non-reducing ends. Starch
15 that is mainly composed of amylose, contains much fewer branches and thus the biosynthetic capacity is reduced. In order to enhance starch biosynthesis when there is no amylopectin production, expression of genes as described in the present invention, could for example form new primers that can replace amylopectin as a source for starch biosynthesis capacity and thereby reduce or eliminate the lost capacity for
20 starch synthesis. To further illustrate the situation the degree of branching in ordinary potato starch is approximately 3.1% while in high amylose starch it is 0.3-1.0% depending on amylose content. This decrease of branching and starch content is further associated with an increase in glucose and fructose content.

25 The increased amylose content and thereby solids content is also advantageous for the processing properties in various applications such as for french fries, potato crisps and other potato based products. In addition to an increased solid content, the inserted genes SEQ ID NO 1 or 3 of the present invention result in the transformation of excess sugars into α -1,4-glucan chains and thereby reducing browning of fried potato products, Maillard reaction, in which amino acids react with free sugars.
30

Furthermore

- (i) any gene of plant origin with the described activity can be used for increasing amylose content and solids
- (ii) the genes can be controlled by any regulating promoter element functional in plant.
- (iii) any starch producing crop of any variety can be transformed with the described genes.
- (iiii) any plant transformation method can be used.
- (iiiii) any binary vector can be used for the insertion of the described genes.
- (iiiiii) the described genes can be combined with any other desired transgenically inserted traits.

The invention further relates to a method for producing amylose by culturing plants which have, compared to a wild type or a genetically modified plant producing already amylose type starch, an increased amylose biosynthesis activity, said proteins comprising the amino acid sequence SEQ ID NO 2 or 4 or a sequence which is derived from one of these sequences by substitution, insertion or deletion of amino acids and which is at least 50% identical at the amino acid level to the sequence SEQ ID NO 2 or 4.

Increased amylose biosynthesis activity compared to the wild type or transgenic line means that the amount of amylose formed is increased by the starch biosynthesis enhancing protein in comparison with the wild type or transgenic line.

This increase in starch or amylose biosynthesis activity is preferably at least 5%, further preferably at least 10%, further preferably at least 20%, further preferably at least 50%, more preferably at least 100%, still more preferably at least 200%, in particular at least 500%, of the protein activity of the wild type or transgenic line.

A "wild type" means the corresponding genetically unmodified starting plant. This plant is preferably *Solanum tuberosum*.

Depending on the context, the term "plant" means a wild type starting plant or a genetically modified starting plant.

"Transgenic plant" or "genetically modified plant" means that the plant contains an additional inserted gene segment that may be foreign or endogenous to the plant species, additional genes or additional gene fragments in sense and/or antisense orientation to a suitable promoter corresponding to the following polypeptides and showing enzymatic activity of a starch branching enzyme I, a starch branching enzyme II and/or

the starch biosynthesis enhancing protein as specified in SEQ ID NO 1 or 3 or polynucleotides having at least 60 % sequence identity thereof.

5 "Amylose type starch" means that the amylose content of the starch is increased compared to the amylose content of starch produced by wild type plants especially wild type potato plants.

10 The starch or amylose biosynthesis activity may be increased in various ways, for example by eliminating inhibiting regulatory mechanisms at the translation and protein levels or by increasing the gene expression of a nucleic acid encoding a starch biosynthesis enhancing protein compared to the wild type or transgenic plant, for example by inducing a gene encoding the starch biosynthesis enhancing protein via activators or by introducing into the plant nucleic acids encoding a starch biosynthesis enhancing protein.

15 According to the invention, increasing the gene expression of a nucleic acid encoding a starch biosynthesis enhancing protein could also mean manipulating the expression of the endogenous starch biosynthesis enhancing protein intrinsic to the plant, in particular in potato plants. This may be achieved, for example, by modifying the promoter
20 DNA sequence of genes encoding a starch biosynthesis enhancing protein. Such a modification which leads to a modified or preferably increased rate of expression of at least one endogenous gene encoding a starch biosynthesis enhancing protein may be carried out by deleting or inserting DNA sequences.

25 It is also possible to modify expression of one or more endogenous starch biosynthesis enhancing protein by applying exogenous stimuli. This may be carried out by particular physiological conditions, i.e. by applying foreign substances.

30 Furthermore, it is possible to achieve a modified or increased expression of at least one endogenous gene encoding a starch biosynthesis enhancing protein by the interaction of a regulatory protein which is modified or is not present in the untransformed plant.

35 In a preferred embodiment, the starch biosynthesis enhancing protein activity is increased compared to the wild type or transgenic plant by increasing the gene expression of a nucleic acid encoding a starch biosynthesis enhancing protein, said starch biosynthesis enhancing protein comprising the amino acid sequence SEQ ID NO 2 or 4 or a sequence which is derived from said sequences by substitution, insertion or deletion of amino acids and which is at least 50% identical at the amino acid level to
40 the sequence SEQ ID NO 2 or 4.

In the case of genomic nucleic acid sequences encoding a starch biosynthesis enhancing protein from eukaryotic sources, which contain introns, preferably already processed nucleic acid sequences such as the corresponding cDNAs are to be used, if the host organism is unable to or cannot be enabled to express the corresponding starch biosynthesis enhancing protein.

In this preferred embodiment, the transgenic plant of the invention thus contains, compared to the wild type or transgenic plant, at least one further gene encoding a starch biosynthesis enhancing protein. In this preferred embodiment, the genetically modified plant of the invention has accordingly at least one transgenic endogenous or exogenous nucleic acid encoding a starch biosynthesis enhancing protein.

Suitable and preferred nucleic acids are described above. In a particularly preferred embodiment, a nucleic acid comprising the sequence SEQ ID NO 1 or 3 is introduced into the plant.

According to the invention, organisms means preferably eukaryotic organisms, such as, for example, yeasts, algae, mosses, fungi or plants, which are capable of producing starch or amylose, either as wild type or enabled by genetic modification. Preferred organisms are photosynthetically active organisms such as, for example, plants which, even as a wild type, are capable of producing starch or amylose type starch.

Particularly preferred organisms are potato plants.

The present invention furthermore relates to the use of proteins comprising the amino acid sequence SEQ ID NO 2 or 4 or a sequence which is derived from this sequence by substitution, insertion or deletion of amino acids and which is at least 50% identical at the amino acid level to the sequence SEQ ID NO 2 or 4 and having starch biosynthesis enhancing activity.

The present invention further relates to the use of nucleic acids SEQ ID NO 1 or 3 or one of the SEQ ID NOs 5, 7, 9, 11 or 13 encoding proteins having a starch biosynthesis enhancing activity in plants.

The transgenic organisms, in particular plants, are preferably prepared by transforming the starting organisms, in particular plants, with a nucleic acid construct containing the above-described nucleic acid, encoding a starch biosynthesis enhancing protein which is functionally linked to one or more regulatory signals ensuring transcription and translation in said organisms.

These nucleic acid constructs in which the coding nucleic acid sequence is functionally linked to one or more regulatory signals ensuring transcription and translation in organisms, in particular in plants, are also referred to as expression cassettes herein below.

5

Accordingly, the invention further relates to nucleic acid constructs, in particular to nucleic acid constructs functioning as expression cassette, which comprise a nucleic acid encoding a starch biosynthesis enhancing protein which is functionally linked to one or more regulatory signals ensuring transcription and translation in organisms, in particular in plants.

10

The regulatory signals preferably comprise one or more promoters ensuring transcription and translation in organisms, in particular in plants.

15 The expression cassettes include regulatory signals, i.e. regulatory nucleic acid sequences, which control expression of the coding sequence in the host cell. According to a preferred embodiment, an expression cassette comprises upstream, i.e. at the 5' end of the coding sequence, a promoter and downstream, i.e. at the 3' end, a polyadenylation signal and, where appropriate, further regulatory elements which are operatively linked to the coding sequence for at least one of the above-described genes
20 located in between. Operative linkage means the sequential arrangement of promoter, coding sequence, terminator and, where appropriate, further regulatory elements in such a way that each of the regulatory elements can properly carry out its function in the expression of the coding sequence.

25

When the organism used is a plant, the nucleic acid constructs and expression cassettes of the invention preferably contain a nucleic acid encoding a plastid transit peptide ensuring localisation in plastids.

30 The preferred nucleic acid constructs, expression cassettes and vectors for plants and methods for preparing transgenic plants and also the transgenic plants themselves are described in examples 2 to 6 below.

The sequences preferred for operative linkage, but not limited thereto, are targeting
35 sequences for ensuring subcellular localisation to plastids such as amyloplasts or chloroplasts but could also mean in the apoplasts, in the vacuole, in the mitochondrion, in the endoplasmic reticulum (ER), in the nucleus, in elaioplasts or in other compartments and translation enhancers such as the tobacco mosaic virus 5'-leader sequence (Gallie et al., Nucl. Acids Res. 15 (1987), 8693-8711).

40

A suitable promoter of the expression cassette is in principle any promoter which is able to control the expression of foreign genes in plants.

5 "Constitutive" promoter means those promoters which ensure expression in numerous, preferably all, tissues over a relatively long period of plant development, preferably during the entire plant development.

Preference is given to using, in particular, a promoter from plants or a promoter originating from a plant virus. Preference is in particular given to the promoter of the 35S transcript of the CaMV cauliflower mosaic virus (Franck et al. (1980) Cell 21:285-294; Odell et al. (1985) Nature 313:810-812; Shewmaker et al. (1985) Virology 140:281-288; Gardner et al. (1986) Plant Mol Biol 6:221-228) or the 19S CaMV promoter (US 5,352,605; WO 84/02913; Benfey et al. (1989) EMBO J 8:2195-2202).

15 Another suitable constitutive promoter is the Rubisco small subunit (SSU) promoter (US 4,962,028), the leguminB promoter (GenBank Acc. No. X03677), the Agrobacterium nopaline synthase promoter, the TR double promoter, the agrobacterium OCS (octopine synthase) promoter, the ubiquitin promoter (Holtorf S et al. (1995) Plant Mol Biol 29:637-649), the ubiquitin 1 promoter (Christensen et al. (1992) Plant Mol Biol 20 18:675-689; Bruce et al. (1989) Proc Natl Acad Sci USA 86:9692-9696), the Smas promoter, the cinnamyl alcohol dehydrogenase promoter (US 5,683,439), the promoters of the vacuolar ATPase subunits or the promoter of a proline-rich wheat protein (WO 91/13991), the Pnit promoter (Y07648.L, Hillebrand et al. (1998), Plant. Mol. Biol. 36, 89-99, Hillebrand et al. (1996), Gene, 170, 197-200) and other promoters of genes 25 whose constitutive expression in plants is known to the skilled worker.

The expression cassettes may also contain a chemically inducible promoter (review: Gatz et al. (1997) Annu Rev Plant Physiol Plant Mol Biol 48:89-108) which may be used to control expression of the starch biosynthesis enhancing protein gene in the 30 plants at a particular time. Promoters of this kind, such as, for example, the PRP1 promoter (Ward et al. (1993) Plant Mol Biol 22:361-366), salicylic acid-inducible promoter (WO 95/19443), a benzenesulfonamide-inducible promoter (EP 0 388 186), a tetracycline-inducible promoter (Gatz et al. (1992) Plant J 2:397-404), an abscisic acid-inducible promoter (EP 0 335 528) and an ethanol- or cyclohexanone-inducible promoter 35 (WO 93/21334), may likewise be used.

Further examples of suitable promoters are fruit ripening-specific promoters such as, for example, the fruit ripening-specific promoter from tomato (WO 94/21794, EP 409 625). Development-dependent promoters partly include the tissue-specific promoters, 40 since individual tissues are naturally formed in a development-dependent manner.

Furthermore, preference is given in particular to those promoters which ensure expression in tissues or parts of the plant, in which, for example, biosynthesis of starch or amylose or of the precursors thereof takes place. Preference is given, for example, to promoters with specificities for leaves, stems, roots, seeds and tubers.

Seed-specific promoters are, for example, the phaseoline promoter (US 5,504,200; Bustos MM et al. (1989) *Plant Cell* 1(9):839-53), the promoter of the 2S albumin gene (Joseffson LG et al. (1987) *J Biol Chem* 262:12196-12201), the legumin promoter (Shirsat A et al. (1989) *Mol Gen Genet* 215(2): 326-331), the USP (unknown seed protein) promoter (Bäumlein H et al. (1991) *Mol Gen Genet* 225(3):459-67), the promoter of the napin gene (US 5,608,152; Stalberg K et al. (1996) *L Planta* 199:515-519), the sucrose-binding protein promoter (WO 00/26388) and the legumin B4 promoter (LeB4; Bäumlein H et al. (1991) *Mol Gen Genet* 225: 121-128; Bäumlein et al. (1992) *Plant Journal* 2(2):233-9; Fiedler U et al. (1995) *Biotechnology (NY)* 13(10):1090f), the *Arabidopsis* oleosin promoter (WO 98/45461), the *Brassica* Bce4 promoter (WO 91/13980) and the vicillin promoter (Weschke et al. 1988, *Biochem. Physiol. Pflanzen* 183, 233-242; Bäumlein H et al. (1991) *Mol Gen Genet* 225(3):459-67).

Further suitable seed-specific promoters are those of the genes coding for high molecular weight glutenine (HMWG), gliadin, branching enzyme, ADP glucose pyrophosphatase (AGPase) and starch synthase. Preference is further given to promoters which allow seed-specific expression in monocotyledons such as e.g. corn, barley, wheat, rye, rice, etc. It is also possible to use advantageously the promoter of the *lpt2* or *lpt1* gene (WO 95/15389, WO 95/23230) or the promoters described in WO 99/16890 (promoters of the hordein gene, the glutelin gene, the oryzin gene, the prolamin gene, the gliadin gene, the glutelin gene, the zein gene, the kasirin gene and the secalin gene).

Examples of tuber-, storage root- or root-specific promoters are the patatin promoter class I (B33), the potato cathepsin D inhibitor promoter and the potato granular bound starch synthase (GBSS) promoter as described in EP-A 0 921 191.

Examples of leaf-specific promoters are the cytosolic FBPase promoter from potato (WO 97/05900), the rubisco (ribulose-1,5-bisphosphate carboxylate) SSU (small sub-unit) promoter and the potato ST-LSI promoter (Stockhaus et al. (1989) *EMBO J* 8:2445-2451).

Further promoters suitable for expression in plants have been described (Rogers et al. (1987) Meth in Enzymol 153:253-277; Schardl et al. (1987) Gene 61:1-11; Berger et al. (1989) Proc Natl Acad Sci USA 86:8402-8406).

5 The site of starch and amylose biosynthesis in potato plants is the amyloplast. Therefore amyloplast-specific targeting and activity of the gene products of the inventive nucleic acids SEQ ID NO 1 or 3 encoding a starch biosynthesis enhancing protein is desirable.

10 The expression may also take place in a tissue-specific manner in all parts of the plant.

A further preferred embodiment therefore relates to a tuber-specific expression of the nucleic acids SEQ ID NO 1 or 3.

15 In addition, a constitutive expression of the gene encoding a starch biosynthesis enhancing protein is advantageous. On the other hand, however, an inducible expression of this gene may also be desirable.

20 An expression cassette is preferably prepared by fusing a suitable promoter to an above-described nucleic acid encoding a starch biosynthesis enhancing protein and, preferably, to a nucleic acid which has been inserted between promoter and nucleic acid sequence and which codes for an amyloplast-specific transit peptide and also to a polyadenylation signal according to familiar recombination and cloning techniques as described, for example, in T. Maniatis, E.F. Fritsch and J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
25 (1989) and in T.J. Silhavy, M.L. Berman and L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience (1987).

30 Particular preference is given to inserted nucleic acid sequences which ensure targeting in the amyloplasts.

35 It is also possible to use an expression cassette in which the nucleic acid sequence encodes a starch biosynthesis enhancing protein fusion protein, one part of the fusion protein being a transit peptide which controls translocation of the polypeptide. Preference is given to amyloplast-specific transit peptides which, after translocation of starch biosynthesis enhancing protein into the amyloplasts, are enzymatically cleaved off the starch biosynthesis enhancing protein part.

40

Particular preference is given to the transit peptide which is derived from the *Nicotiana tabacum* plastid transketolase or from another transit peptide (e.g. the transit peptide of the rubisco small subunit or of ferredoxin NADP oxidoreductase and also of isopentenyl pyrophosphate isomerase-2) or from its functional equivalent.

5

Further examples of a plastid transit peptide are the transit peptide of the plastid isopentenyl pyrophosphate isomerase-2 (IPP-2) from *Arabidopsis thaliana* and the transit peptide of the ribulose biphosphate carboxylase small subunit (rbcS) from pea (Guerineau, F, Woolston, S, Brooks, L, Mullineaux, P (1988) An expression cassette for targeting foreign proteins into the chloroplasts. Nucl. Acids Res. 16: 11380).

10

Plant genes of the invention which encode a plant starch biosynthesis enhancing protein may already contain the nucleic acid sequence which encodes a plastid transit peptide. In this case, a further transit peptide is not required. For example, the *Solanum tuberosum* sequences of the starch biosynthesis enhancing protein of the invention SEQ ID NO 1 or 3 contain already a transit peptide sequence.

15

The nucleic acids of the invention may be prepared synthetically or obtained naturally or comprise a mixture of synthetic and natural nucleic acid components and may also be composed of various heterologous gene sections of various organisms.

20

As described above, preference is given to synthetic nucleotide sequences with codons which are preferred by plants. These codons which are preferred by plants may be determined from codons which have the highest frequency in proteins and which are expressed in most of the interesting plant species.

25

When preparing an expression cassette, it is possible to manipulate various DNA fragments in order to obtain a nucleotide sequence which expediently can be read in the correct direction and is provided with a correct reading frame. The DNA fragments may be linked to one another by attaching adaptors or linkers to said fragments.

30

It is furthermore possible to use manipulations which provide appropriate restriction cleavage sites or which remove excess DNA or restriction cleavage sites. In those cases for which insertions, deletions or substitutions such as, for example, transitions and transversions are suitable, in vitro mutagenesis, primer repair, restriction or ligation can be used.

35

Preferred polyadenylation signals are polyadenylation signals functional in plants, exemplified by those which correspond essentially to T-DNA polyadenylation signals from *Agrobacterium tumefaciens*, in particular of the T-DNA gene 3 (octopine synthase)

40

or OCS terminator, the complete sequence of the Ti plasmid pTiACH5 (Gielen et al., EMBO J. 3, 835 -846(1984) or functional equivalents.

5 The invention further relates to the use of the nucleic acids SEQ ID NO 1 or 3 for increasing the starch or amylose content in plants, e.g. potato plants which, as wild type, are capable of producing starch or amylose, see examples 2 to 13.

10 The invention is not limited to the over-expression of the nucleic acid sequences SEQ ID NO 1 or SEQ ID NO 3 in plants especially potato plants.

15 The over-expression of both nucleic acid sequences SEQ ID NO 1 and 3 in a plant can be used for enhancing amylose biosynthesis, see examples 14-16. Constructs containing the nucleic acids SEQ ID NO 1 and SEQ ID NO 3 can also be used for increasing the starch content or the amylopectin content in plants. These constructs can be made on the same T-DNA driven by one promoter each. These constructs can also be made on the same T-DNA in tandem driven by the same promoter. These constructs can also be transformed using more than one construct, either at the same time (co-transformation) or in different transformation events.

20 The above-described proteins and nucleic acids may be used for producing starch or amylose in transgenic plants.

25 The transfer of foreign genes into the genome of an organism, in particular of a plant, is referred to as transformation.

For this purpose, methods known per se for transforming plants and regenerating plants from plant tissues or plant cells can be used, in particular in plants, for transient or stable transformation, e.g. as described in example 2.

30 Suitable methods for the transformation of plants are the protoplast transformation by polyethylene glycol-induced DNA uptake, the biolistic method using the gene gun - also known as particle bombardment method, electroporation, the incubation of dry embryos in a DNA-containing solution, microinjection and the above-described Agrobacterium-mediated gene transfer. Said methods are described, for example, in B. Jenes et al.,
35 Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press (1993), 128-143 and in Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225).

The construct to be expressed is preferably cloned into a vector which is suitable for transforming *Agrobacterium tumefaciens*, for example pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984), 8711) or preferably pSUN2 (WO 02/00900).

- 5 Accordingly, the invention furthermore relates to vectors containing the above-described nucleic acids, nucleic acid constructs or expression cassettes.

Agrobacteria which have been transformed with an expression cassette can be used in a known manner for the transformation of plants, for example by bathing injured leaves
10 or leaf sections in an *Agrobacterium* solution and then culturing them in suitable media.

Apart from in plants, the expression cassette may also be used for transforming bacteria, in particular cyanobacteria, mosses, yeasts, filamentous fungi and algae.

- 15 Genetically modified plants, also referred to as transgenic plants herein below, are preferably prepared by cloning the fused expression cassette which expresses a starch biosynthesis enhancing protein into a vector, for example pBin19, which is suitable for transforming *Agrobacterium tumefaciens*.

20 *Agrobacteria* which have been transformed with such a vector may then be used in a known manner for the transformation of plants, in particular of crop plants, for example by bathing injured leaves or leaf sections in an *Agrobacterium* solution and then culturing them in suitable media.

- 25 The transformation of plants by *Agrobacteria* is described, inter alia, in F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press, 1993, pp. 15-38. Transgenic plants which contain a gene for expression of a nucleic acid encoding a starch biosynthesis enhancing protein, which has been integrated into the expression
30 cassette, can be regenerated in a known manner from the transformed cells of the injured leaves or leaf sections.

A host plant is transformed with a nucleic acid SEQ ID NO 1 or 3 encoding a starch biosynthesis enhancing protein by incorporating an expression cassette as insertion
35 into a recombinant vector whose vector DNA comprises additional functional regulatory signals, for example sequences for replication or integration. Suitable vectors are described inter alia, in Methods in Plant Molecular Biology and Biotechnology (CRC Press), chapter 6/7, pp. 71-119 (1993).

By way of example, the plant expression cassette may be incorporated into a derivative of the transformation vector pBin-19 with 35s promoter (Bevan, M., Nucleic Acids Research 12: 8711-8721 (1984).

5 Using the above-cited recombination and cloning techniques, it is possible to clone the expression cassettes into suitable vectors for maintenance and propagation of genetic material for example in *E. coli*. Suitable cloning vectors are, inter alia, pBR322, pUC series, M13mp series, pBluescript and pACYC184. Particularly suitable are binary vectors which can replicate both in *E. coli* and in agrobacteria.

10

The invention therefore further relates to the use of the above-described nucleic acids or of the above-described nucleic acid constructs, in particular of the expression cassettes, for preparing genetically modified plants or for transforming plants, plant cells, plant tissues or parts of plants.

15

The use is preferably aimed at increasing the starch or amylose content of the plant, of the tubers or in other parts of the plant.

20

The use is most preferably aimed at increasing the starch or amylose content of wild-type or transgenic potato plants and especially the tubers of wild-type or transgenic potato plants.

25

Accordingly, the invention further relates to a method for preparing genetically modified plants by introducing an above-described nucleic acid or an above-described nucleic acid construct into the genome of the starting organism.

30

The invention further relates to the genetically modified organisms, the genetic modification increasing the activity of a starch biosynthesis enhancing protein compared to a wild type or transgenic plant and the starch biosynthesis enhancing protein comprising the amino acid sequence SEQ ID NO 2 or 4 or a sequence which is derived from this sequence by substitution, insertion or deletion of amino acids and which is at least 50% identical at the amino acid level to the sequence SEQ ID NO 2 or 4.

35

As illustrated above, the starch biosynthesis enhancing protein activity is increased compared to the wild type or transgenic plant preferably by increasing the gene expression of a nucleic acid encoding a starch biosynthesis enhancing protein.

40

In a further preferred embodiment, gene expression of a nucleic acid encoding a starch biosynthesis enhancing protein is increased, as illustrated above, by introducing nucleic acids encoding a starch biosynthesis enhancing protein into the organism and

thus by over-expressing nucleic acids encoding a starch biosynthesis enhancing protein.

Such transgenic plants, their propagation material and their plant cells, plant tissues,
5 plant parts or tubers are a further subject of the present invention.

Genetically modified plants of the invention, which have an increased starch or amylose content and which can be consumed by humans and animals, can also be used as food- or feedstuffs or as feed and food supplements, for example directly or after processing known per se. The genetically modified plants may furthermore be used for
10 producing starch or amylose-containing extracts of said plant and/or for producing feed and food supplements.

The invention further relates to:

- 15
- I. A polynucleotide that encodes a polypeptide of SEQ ID NO 1 or 3.
 - II. A polynucleotide comprising at least 30 contiguous bases of SEQ ID NO 1 or 3.
 - 20 III. A polynucleotide having at least 60 % sequence identity to SEQ ID NO 1 or 3, wherein the identity is based on the entire coding sequence.
 - IV. A polynucleotide having at least 60 % sequence identity to SEQ ID NO 1 or 3, wherein the % sequence identity is based on the entire sequence.
 - 25 V. A polynucleotide which selectively hybridizes, under stringent conditions and a wash in 2 X SSC at 50 °C, to a hybridization probe derivable from the polynucleotide sequence as set forth in SEQ ID NO 1 or 3, or from the genomic sequence.
 - 30 VI. A polynucleotide complementary to a polynucleotide of V.
 - VII. The polynucleotide of I, wherein the starch or amylose biosynthesis enhancing polynucleotide is from *Solanum tuberosum*.
 - 35 VIII. The polynucleotide of I encoding a polypeptide, which after over-expression in a plant cell increases the starch or amylose content.
 - IX. The polynucleotide of I in antisense orientation, which after expression in a
40 plant cell decreases the starch or amylose content.

- X. A vector comprising at least one polynucleotide of I.
- 5 XI. An expression cassette comprising at least one polynucleotide of I operably linked to a promoter, wherein the polynucleotide is in sense or antisense orientation.
- XII. A host cell which is introduced with at least one expression cassette of X.
- 10 XIII. The host cell of XI that is a plant cell.
- XIV. A transgenic plant comprising at least one expression cassette of XI.
- XV. The transgenic plant of XIII, wherein the plant is *Solanum tuberosum*.
- 15 XVI. A tuber from the transgenic plant of XIV.
- XVII. An isolated protein comprising a member selected from the group consisting of:
- 20 a) a polypeptide comprising at least 10 contiguous amino acids of SEQ ID NO 2 or 4,
- b) a polypeptide which is a plant starch biosynthesis enhancing protein,
- c) a polypeptide comprising at least 55 % sequence identity to SEQ ID NO 2 or 4, wherein the sequence identity is based on the entire sequence and has at least one epitope in common with a starch biosynthesis enhancing protein.
- 25 d) a polypeptide encoded by a polynucleotide selected from SEQ ID NO 1 or 3,
- e) a polypeptide of SEQ ID NO 2 or 4.
- 30 XVIII. The protein of XVII, wherein the polypeptide is catalytically active.
- XIX. A ribonucleic acid sequence encoding the protein of XVIII.
- XX. A method for modulating the level of starch biosynthesis enhancing protein in a plant, comprising:
- 35 a) stably transforming a plant cell with a polynucleotide coding for a starch biosynthesis enhancing protein operably linked to a promoter, wherein the polynucleotide is in sense or antisense orientation

- b) growing the plant cell under plant growing conditions to produce a regenerated plant capable of expressing the polynucleotide for a time sufficient to modulate the level of starch biosynthesis enhancing protein in the plant.
- 5 XXI. The method of XX, wherein the polynucleotide coding for a starch biosynthesis enhancing protein is selected from SEQ-ID NO 1 or 3.
- XXII. The method of XX, wherein the plant is *Solanum tuberosum*.
- 10 XXIII. The method of XX, wherein activity of the starch biosynthesis enhancing protein is increased.
- XXIV. A method for modulating the level of starch or amylose in a plant, comprising:
- 15 a) stably transforming a plant cell with a polynucleotide coding for a starch biosynthesis enhancing protein operably linked to a promoter, wherein the polynucleotide is in sense or anti-sense orientation,
- b) growing the plant cell under plant growing conditions to produce a regenerated plant capable of expressing the polynucleotide for a time sufficient to modulate level of starch or amylose in the plant.
- 20 XXV. A method for modulating the level of starch or amylose in a plant, comprising:
- a) stably transforming a plant cell with a polynucleotide encoding a starch biosynthesis enhancing protein operably linked to a promoter, wherein the polynucleotide is in sense or anti-sense orientation.
- 25 b) growing the plant cell under plant growing conditions to produce a regenerated plant capable of expressing the polynucleotide for a time sufficient to modulate level of starch or amylose in the plant.
- 30 XXVI. The method of XXIV wherein the polynucleotide coding for a starch biosynthesis enhancing protein is selected from SEQ ID NO 1 or 3.

Some of the terms used further on in the specification are defined at this point.

35 "Enzymatic activity/activity assay": the term enzymatic activity describes the ability of an enzyme to convert a substrate into a product. In this context, both the natural substrate of the enzyme and a synthetic modified analog of the natural substrate can be used. The enzymatic activity can be determined in what is known as an activity assay via the increase in the product, the decrease in the starting material, the decrease or increase in a specific cofactor, or a combination of at least two of the aforementioned

40 parameters as a function of a defined period of time.

"Functional equivalents" in the present context describe nucleic acid sequences which hybridize under standard conditions with the nucleic acid sequence encoding the starch biosynthesis enhancing protein or portions of the nucleic acid sequence encoding the starch biosynthesis enhancing protein, and which are capable of bringing about the expression of an enzymatically active plant starch biosynthesis enhancing protein in a cell or an organism.

It is advantageous to use short oligonucleotides of a length between 10 to 50bp, preferably 15-40bp, for example of the conserved or other regions, which can be determined via comparisons with other related genes in a manner known to the skilled worker for the hybridization. Alternatively, it is also possible to use longer fragments of the nucleic acids according to the invention or the complete sequences for the hybridization. These standard conditions vary depending on the nucleic acid used, namely oligonucleotide, longer fragment or complete sequence, or depending on which type of nucleic acid, that is DNA or RNA, is being used for the hybridization. Thus, for example, the melting temperatures for DNA:DNA hybrids are approx. 10°C lower than those of DNA:RNA hybrids of equal length. Suitable hybridization conditions are described above.

A functional equivalent is furthermore also understood as meaning, in particular, natural or artificial mutations of the relevant nucleic acid sequences of the plant starch biosynthesis enhancing protein and their homologs from other organisms which make possible the expression of the enzymatically active plant starch biosynthesis enhancing protein in a cell or an organism.

Thus, the scope of the present invention also extends to, for example, those nucleotide sequences which are obtained by modification of the nucleic acid sequence of a starch biosynthesis enhancing protein. The purpose of such a modification can be, for example, the insertion of further cleavage sites for restriction enzymes, the removal of excess DNA, or the addition of further sequences. Proteins which are encoded via said nucleic acid sequences should still maintain the desired functions, despite the deviating nucleic acid sequence.

The term functional equivalent may also refer to the protein encoded by the nucleic acid sequence in question. In this case, the term functional equivalent describes a protein whose amino acid sequence is up to a specific percentage identical with that of the starch biosynthesis enhancing protein.

Functional equivalents thus encompass naturally occurring variants of the sequences described herein, and also artificial, for example chemically synthesized, nucleic acid

sequences adapted to the codon usage, or the amino acid sequences derived therefrom.

5 In general, it can be said that functional equivalents independently of the amino acid sequence in question (encoded by a corresponding nucleic acid sequence) have in each case the enzymatic activity of a starch biosynthesis enhancing protein.

"Reporter genes" encode readily quantifiable proteins. Using these genes, an assessment of transformation efficacy or of the site or time of expression can be made via
10 growth, fluorescence, chemoluminescence, bioluminescence or resistance assay or via photometric measurement (intrinsic color) or enzyme activity. Very especially preferred in this context are reporter proteins (Schenborn E, Groskreutz D. Mol. Biotechnol. 1999; 13(1):29-44) such as the "green fluorescence protein" (GFP) (Gerdes HH and Kaether C, FEBS Lett. 1996; 389(1):44-47; Chui WL et al., Curr. Biol. 1996, 6:325-330;
15 Leffel SM et al., Biotechniques. 23(5):912-8, 1997), chloramphenicol acetyl transferase, a luciferase (Giacomin, Plant Sci. 1996, 116:59-72; Scikantha, J. Bact. 1996, 178:121; Millar et al., Plant Mol. Biol. Rep. 1992 10:324-414), and luciferase genes, in general b-galactosidase or b-glucuronidase (Jefferson et al., EMBO J. 1987, 6, 3901-3907), the Ura3 gene, the Ilv2 gene, the 2-desoxyglucose-6-phosphate phosphatase gene,
20 b-lactamase gene, the neomycin phosphotransferase gene, the hygromycin phosphotransferase gene, or the BASTA (= gluphosinate) resistance gene.

"Significant increase": referring to the enzymatic activity, is understood as meaning the increase in the enzymatic activity of the enzyme incubated with a candidate compound
25 in comparison with the activity of an enzyme not incubated with the candidate compound, which lies outside an error in measurement.

"Substrate": Substrate is the compound which is recognized by the enzyme in its original function and which is converted into a product by means of a reaction catalyzed by the enzyme.
30

Preferably, the plant starch biosynthesis enhancing protein is encoded by a nucleic acid sequence comprising

- 35 a) a nucleic acid sequence shown in SEQ ID NO 1 or 3; or
- b) a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO 2 or 4 by back translation; or

- c) a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from a functional equivalent of the amino acid sequence shown in SEQ ID NO 2 or 4, which has an identity with SEQ ID NO 2 or 4 of at least 50%, by back translation.

5

The functional equivalent of SEQ ID NO 2 or 4 set forth in c) has an identity of at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57% preferably at least 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, and 70% more preferably 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85% most preferably at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identity with the SEQ ID NO 2 or 4.

10

Potato varieties used for starch production as well as genotypes with a high amylose content are transformed with gene constructs as described and in example 17 for the over-expression of a starch biosynthesis enhancing protein. The over-expression of StGH1 or StGH2 in potato plants will result in an increased starch or amylose content of the transgenic plant compared to the starting plant. The increase of starch content in the transgenic lines can be seen in table 9. The lines also show an increased harvest weight when grown in the greenhouse (table 6) thus resulting in an increased starch yield.

20

Example 1

Complementation study in yeast

Yeast contains two self-glycosylating proteins, Glg1p and Glg2p, which yield primers for the initiation of glycogen synthesis. For glycogen synthesis to take place in yeast it is required that either gene is functional. Yeast strain CC9, contain knock-out mutations for both genes and is therefore a null mutant regarding this specific biosynthetic function and is therefore unable to produce glycogen (Cheng, C. et al., Molecular and Cellular Biology (1995), 6632-6640). CC9 was used as a basis for complementation experiments with the isolated potato genes in order to validate their function by restoring glycogen biosynthesis in the CC9 strain. The potato genes were cloned in a yeast plasmid, pRS414 (Stratagene), and expressed with various yeast controlling elements such as Gal1, Adh1 and Glg2p promoters. CC9 was transformed by the resulting plasmids using LiCl and electroporation (Multiporator, Eppendorf). Transformed yeast colonies growing on appropriate media plates were screened by immersing in iodine solution. Wild type yeast producing glycogen is stained red brown by iodine while the null mutant CC9 is not stained. CC9 expressing the potato genes, StGH1 and StGH2, will stain red brown, when the isolated genes complement a glycogenin function in yeast and thus carry the desired function.

40

Example 2

Transformation method

- 5 Fully expanded leaves from in vitro propagated potato plants are diagonally cut in 2-4 pieces and precultivated on MC-plates for 2-3 days at 23-24°C.

- Agrobacterium tumefaciens strain LBA4404 containing pHS1, pHS2, pHS3 pHS4, pHASHS2, pHASHS4, pHASHS5, pHASHS6, pHASHS7 or pHASHS8 are grown in
 10 YEB medium with 100µg rifampicin and 25µg/ml kanamycin over night on constant shaking (200 rpm) at 28°C.

- The Agrobacterium culture is prepared for infection by dilution 1:20 with MS10 medium. The leaf explants are infected for 8-10 min in the bacterial solution and afterwards
 15 drained on filter paper for 5-20 seconds. The leaf segments are placed on the MS300 plates for 2 days co-cultivation under modest light at 23-24°C. At the end of co-cultivation the leaf segments are moved to M400 plates containing 400 g/l Claforan to suppress bacterial growth. After 4-5 days the explants are moved to selection medium MS400 supplemented with 400 g/l Claforan. For explants transformed with pHS1 and
 20 pHS2 50µM kanamycin was included in the media and for explants transformed with pHS3, pHS4, pHASHS2, pHASHS4, pHASHS5, pHASHS6, pHASHS7 and pHASHS8 0.5 M Imazamox was added to the media.
- Leaf segments are transferred to fresh MS 400 selection medium every fortnight. The regenerated putative transgenic shoots are collected and cultivated on MS30 plates
 25 with 200 g/l Claforan aiming at shoot elongation.

When the shoots are 3-5 cm long, 1-2 cm are cut off and grown on microtuber medium in the dark at 25°C. After 2-5 weeks microtubers are produced.

MC plates	MS300
MS300 plates with 1.5-2 ml liquid MS100 medium and covered with one sterile filter paper	4.4 g/l MS-medium 2 mg/l naphthyl acetic acid 1 mg/l 6-benzyl amino pyridine 3% (w/v) sucrose pH 5.2

MS10 4.4 g/l MS-medium (murashige and Skoog) 1% (w/v) sucrose pH 5.8	MS400 4.4 g/l MS-medium 2 g/l zeatine 0.01 mg/l naphthyl acetic acid 0.1 mg/l gibberellic acid 10% (w/v) sucrose 400 mg/l claforan 0.5 μ M Imazamox or 50 μ M kanamycin pH .8
MS30 4.4 g/l MS-medium 3% (w/v) sucrose pH 5.8	Microtuber medium 4.4 g/l MS-medium 2.5 mg/l kinetin 0.5 mg/l abscisic acid 8% sucrose 200 mg/cloforan
MS100 4.4 mg/l MS-medium 30 /l sucrose 0.5 m/lg thiamin-HCl 0.5 mg/l pyridoxin-HCl 1 mg/l nicotinacid 0.5 mg/l kinetin 29.8 mg/l ferrous sulfate hepta hydrate 1 mg/l 2,4-Dichlorophenoxyacetic acid 2 g/l caseinhydrolysate pH 5.2	

Example 3

Transgenic plant AM 99-2003

- 5 High amylose potato lines can be produced for example by using antisense, RNAi or antibody technology that target the two starch branching enzymes starch branching enzyme 1 (SBE1) and starch branching enzyme 2 (SBE2).

- 10 The high amylose potato line AM99-2003 is produced by inhibition of the starch branching enzyme activities in the parental line Dinamo. Transformation is made with a construct of SBE1 and SBE2 in antisense orientation driven by the gbss promoter.

pBluescript containing a 1620bp fragment of the 3' end of Sbe1 between EcoRV and SpeI is cut open with SpeI (blunt) and XbaI and ligated with a 1243bp SstI (blunt) and XbaI fragment of the 3' end of Sbe2. The Sbe2 and Sbe1 complex is cut out with
5 EcoRV and XbaI and ligated to the SmaI and XbaI opened up binary vector pHo3.1, see figure 8. The final vector is named pHAbE12A, see figure 9 and nucleic acid sequence SEQ ID NO 15. pHo3.1 is based on pGPTVKan (Becker, D. et al., Plant Molecular Biology 20 (1992), 1195-1197) with the addition of the 987bp gbss promoter cloned at the HindIII site of pGPTVKan and the uidA gene is deleted by SmaI and SstI.

10 The parental line Dinamo is transformed with the construct pHAbE12A as described in example 2.

Example 4

15 Down-regulation of StGH1 and StGH2 genes in potato by antisense

The StGH1 and StGH2 genes were down-regulated in potato by transformation with the genes in antisense direction in relation to a plant regulatory element. The respective antisense genes were cloned in a binary vector driven by a tuber specific gbss
20 promoter. NptII, yielding resistance to the antibiotic kanamycin, was used as selection marker. Two varieties were transformed, Prevalent and Producent. The shoots were selected on 50 µM kanamycin, which is a standard kanamycin concentration used for potato transformation (Ooms, G et al., Theoretical and Applied Genetics 73:744-750 (1987) and Tavazza, R. et al., Plant Science 59 (1988), 175-181).

25 Example 5

Over-expression of StGH1 and StGH2 genes in potato

The StGH1 and StGH2 genes were over-expressed in potato driven by the tuber specific promoter gbss. A mutated AHAS gene was used as selection marker yielding
30 tolerance to the Imazamox herbicides. Two potato varieties were transformed, Desiree and AM99-2003 a transgenic high amylose line with a 40% decrease in starch content compared to its parental line.

35 Example 6

Selection of transgenic lines

Non-transgenic escapes were identified and discarded by a PCR screening method. DNA was extracted according to DNeasy 96 Plant protocol (Qiagen). In a 96 well
40 microtiter plate, 10-15 mg leaf tissue was added to each well together with a 5mm steel

ball, each well then representing one individual shoot. The plates were frozen in N₂(l) before homogenisation. The homogenisation was done at 30Hz in a Mixermill300 for 1 min. The DNA was at the end of the extraction protocol eluted in 75µl H₂O.

- 5 Specific primers for nptII and AHAS were used for the amplification of a 246bp fragment respective a fragment of 509 bp for selection of successfully transformed lines.

Npt2_for 5'-AGCAAGGTGAGATGACAGGAGATC-3'

Npt2_rev 5'-CAGACAATCGGCTGCTCTGATG-3'

10

AHAS1_frw: 5'-AACAACAACATCTTCTTCGATC-3'

AHAS1_rev: 5'-TAACGAGATTTGTAGCTCCG-3'.

The PCR reactions were with the extracted DNA setup and run as follow:

15

Reaction:

	10x PCR Mix	2,0 µl
	Primer frw (25µM)	0,4 µl
20	Primer rev (25µM)	0,4 µl
	dNTPs (10mM)	0,4 µl
	RedTAQ (Sigma)	1,0 µl
	Templat (~20ng/µl)	4,0 µl
	H ₂ O	11,8 µl

25

PCR program:

	94°C	30 s	
	59°C	30 s	x29 cycles
30	72°C	30s	
	72°C	7 min	
	8°C	old	

Example 7

35 Gene expression analysis

The gene expression levels of the StGH1 and StGH2 genes were analysed in the transgenic potato lines with real-time PCR (ABI prism 7900HT, Applied Biosystems). With real-time PCR the change of gene expression can be analysed regarding RNA expression levels. For pHS1 and pHS2 transgenic lines, expression of both sense and

40

antisense RNA of StGH1 and StGH2 was measured, while in pHS3 and pHS4 transgenic lines the change in StGH1 and StGH2 mRNA expression was analysed.

- 5 The target for pHS1 and pHS2 is to reduce transcript levels of StGH1 and StGH2 respectively while the target for pHS3 and pHS4 is to increase transcript levels of the respective genes.

- 10 RNA was isolated from microtubers of the transgenic potato lines and mother varieties using Invisorb Spin Plant-RNA mini kit (Invitex). A reverse transcription reaction was made with 250 ng total RNA in 25 µl total reaction volume using TaqMan reverse transcription reagents (Applied Biosystems). Separate and specific primers (see table 1) were designed and used for the reverse transcription reaction in order to be able to differentiate the endogenous expression from the antisense RNA expression of the respective genes.

15

StGH1 sense RNA	5'-TGAAGACAGCACAAAACCTGG-3'
StGH1 antisense RNA	5'-GTGAAAGTTTGAACGCACAC-3'
StGH2 sense RNA	5'-AGTGCCATAACATGCTTTCC -3'
StGH2 antisense RNA	5'-CACATTTTCAGCTGTTGATGGA-3'

Table 1

- 20 5 µl of the reverse transcription reaction was used in triplicate analyses together with specific sequence detection primers, TaqMan MGB probe (see table 2) and UMM mastermix (Applied Biosystems) and determined with real-time PCR according to the suppliers instructions.

StGH1
Forward Primer: TCGAGTCGCCACGTAGAACTC
Reverse primer: GAAATGCGTATGCGACTATGATG
TaqMan probe: AGTCTCTCGGAGTTCC
StGH2
Forward primer: GGTGCTGATCCTCCAGTTCTCT
Reverse primer: GTCCCTGAAGCATAACCAAGGT
TaqMan probe: TTCTGCACTACTTAGGCCT

25 Table 2

Down-regulation of the two genes resulted in a decrease in gene expression in transgenic lines compared to their mother varieties in the order of 50-95 %.

- 5 Over-expression of the two genes resulted in a 2-10 times increase in gene expression in transgenic lines compared to their mother varieties.

Line No.	Construct	Variety	ΔCt s	Times increase or decrease in gene expression compared to parental line
P01-041-84	pHS1	Producent	-1,14	-1,3
P02-325-1	pHS1	Producent	-2,03	-4,1
P02-325-9	pHS1	Producent	-1,47	-2,2
P02-325-11	pHS1	Producent	-1,25	-1,6
P02-325-15	pHS1	Producent	-2,52	-6,3
P02-325-25	pHS1	Producent	-2,0	-4
P02-325-27	pHS1	Producent	-1,64	-2,7
P02-325-33	pHS1	Producent	-1,59	-2,5
P02-325-34	pHS1	Producent	-1,52	-2,3
P02-325-63	pHS1	Producent	-1,53	-2,3
P02-300-37	pHS2	Prevalent	-1,27	-1,6
P02-300-66	pHS2	Prevalent	-1,04	-1,1
P02-300-71	pHS2	Prevalent	-1,13	-1,3
P02-300-73	pHS2	Prevalent	-1,1	-1,2
P02-300-80	pHS2	Prevalent	-2,12	-4,5
P02-300-127	pHS2	Prevalent	-1,67	-2,8
P02-300-140	pHS2	Prevalent	-3,96	-15,7
P02-303-31	pHS2	Prevalent	-1,16	-1,4
P02-303-64	pHS2	Prevalent	-1,15	-1,3
P02-305-54	pHS2	Prevalent	-1,33	-1,8
P02-320-24	pHS2	Prevalent	-1,03	-1,1
P02-307-4	pHS3	Desirée	1,82	3,3
P02-307-5	pHS3	Desirée	2,68	7,2
P02-307-12	pHS3	Desirée	2,67	7,1
P02-307-14	pHS3	Desirée	1,83	3,3
P02-307-15	pHS3	Desirée	1,79	3,2
P02-307-33	pHS3	Desirée	3,21	10,3
P02-307-43	pHS3	Desirée	2,7	7,3
P02-307-51	pHS3	Desirée	2,73	7,5
P02-307-80	pHS3	Desirée	2,78	7,7
P02-307-87	pHS3	Desirée	1,02	1,1
P02-307-148	pHS3	Desirée	1,88	3,5

Line No.	Construct	Variety	Δ Ct s	Times increase or decrease in gene expression compared to parental line
P02-309-63	pHS3	AM99-2003	1,64	2,7
P02-309-111	pHS3	AM99-2003	1,34	1,8
P02-309-106	pHS3	AM99-2003	1,75	3,1
P02-311-59	pHS3	AM99-2003	1,17	1,4
P02-312-15	pHS4	AM99-2003	1,03	1,1
P02-313-21	pHS4	AM99-2003	1,54	2,4
P02-317-2	pHS4	AM99-2003	1,2	1,4

Table 3: Gene expression analysis based on Real- Time PCR

Example 8

5 Dry matter analysis

Dry matter has been analyzed on microtubers from transgenic lines transformed with pHS1, pHS2, pHS3 and pHS4 showing a down-regulation or over expression of the genes. Since starch normally contribute to more than 80% of the dry matter in potato tubers, an increase or decrease in starch content will affect also the dry matter content.

Two microtubers of each line were harvested when they had reached maturity. Dry matter was calculated for mature microtubers weighed before and after 72 hours drying at 60°C. For comparison microtubers from the varieties Dinamo, Desiree, Prevalent, Producent and P737 with starch contents between 13 and 28% (when grown in field) were used. The starch content of microtubers is not as high as starch content of field grown tubers. However dry matter content can readily be compared in microtubers and that value is well correlated to the determined starch content in field grown tubers. In table 4 the average dry matter for the different varieties, calculated on ten or more microtubers, is shown.

Variety	Starch content field grown tubers	Dry matter microtubers
AM99-2003	13%	14,8
Desirée	16%	16,1
Producent	22%	19,2
Prevalent	22%	19,7
P737	28%	21,6

Table 4: Dry matter content of 5 varieties based on 10 or more microtubers

One of each pHS1 and pHS2 with confirmed decrease in gene-expression have been analyzed for dry matter so far. Those two have a decrease in dry matter of 7 and 11% compared to their mother varieties.

- 5 For the pHS3 lines 8 of 9 of the confirmed over-expressed lines show an increase of up to 36% in dry matter. See table 5.

Line No.	Construct	Variety	Dry matter in relation to parental line (%)
41-84	pHS1	Producent	89
300-127	pHS2	Prevalent	93
300-140	PHS2	Prevalent	96
307-4	pHS3	Desirée	106
307-5	pHS3	Desirée	117
307-15	pHS3	Desirée	124
307-33	pHS3	Desirée	116
307-57	pHS3	Desirée	136
309-63	pHS3	AM99-2003	134
309-106	pHS3	AM99-2003	109
309-111	pHS3	AM99-2003	108

- 10 Table 5: Dry matter content on transgenic lines with confirmed down-regulation or over-expression of the StGH1 and StGH2 genes

Example 9

Starch content analysis

- 15 For analysis of starch content a total starch assay procedure from Megazyme International Ireland Ltd., Bray, Co.Wicklow, Ireland (AOAC Method 996.1; AACC method 76.13; ICC standard method No. 168) was used according to the suppliers instructions.

- 20 Starch content was analysed on microtubers from all transgenic lines transformed with pHS1, pHS2, pHS3 and pHS4. The microtubers were harvested when they had reached maturity. Mature microtubers were ground and maltosaccharides and free glucose residues were washed away with ethanol. The microtuber starch was treated with DMSO to ensure the complete solubilisation of samples with high levels of resistant starch, as the high amylose clones.

25

Samples were analyzed with a standard spectrophotometric assay procedure. The transgenic lines were compared to potato varieties with known starch content ranging

from 8% to 30%. The results give an indication on the change in starch content related to the genetic modification of the different transgenic lines.

Example 10

5 Greenhouse trial

Harvest weight and dry matter content was measured on pHS3 and pHS4 transgenic lines and their mother varieties grown in the greenhouse. The harvest weight from 10 greenhouse grown pots was measured, see table 6.

10

Line No.	Construct	Parental line	Harvest weight (g)	Increase in harvest weight compared to parental line %
AM99-2003			1150	
Desirée			1500	
P02-307-4	pHS3	Desirée	1900	27
P02-307-5	pHS3	Desirée	2050	37
P02-307-12	pHS3	Desirée	2250	50
P02-307-15	pHS3	Desirée	1950	30
P02-307-33	pHS3	Desirée	1950	30
P02-307-43	pHS3	Desirée	2000	33
P02-307-51	pHS3	Desirée	1950	30
P02-307-80	pHS3	Desirée	1750	17
P02-309-63	pHS3	AM99-2003	1550	35
P02-309-114	pHS3	AM99-2003	1250	9
P02-316-25	pHS3	AM99-2003	1450	26
P02-316-111	pHS3	AM99-2003	1300	13
P02-314-4	pHS4	Desirée	1650	10
P02-314-15	pHS4	Desirée	1600	7
P02-314-35	pHS4	Desirée	1850	23
P02-314-40	pHS4	Desirée	1700	13
P02-312-15	pHS4	AM99-2003	1350	17
P02-313-21	pHS4	AM99-2003	1300	13
P02-313-42	pHS4	AM99-2003	1650	43
P02-317-2	pHS4	AM99-2003	1350	17

Table 6: Greenhouse harvest weight of lines over-expressing StGH1 (pHS3) or StGH2 (pHS4).

The harvest weight was increased up to 43% in the transgenic lines compared to their mother varieties. The results show that over-expression of StGH1 and StGH2 results in an increase in total harvest weight.

- 5 The lines grown in the greenhouse were analyzed for dry matter content. Slices of three tubers of each line were dried in a freeze dryer for 72 hours with weighing prior and after the drying. The dry matter results can be seen in table 7 and present a mean value of the three analyses.

Line No.	Construct	Parental line	Mean value	Increase in dry matter in comparison with parental line %
AM99-2003			19,15	
Desiree			19,25	
P02-307-4	pHS3	Desiree	21,29	11
P02-307-5	pHS3	Desiree	19,92	3
P02-307-12	pHS3	Desiree	22,46	17
P02-307-14	pHS3	Desiree	20,45	6
P02-307-15	pHS3	Desiree	22,12	15
P02-307-33	pHS3	Desiree	21,03	9
P02-307-43	pHS3	Desiree	19,68	2
P02-307-80	pHS3	Desiree	19,93	4
P02-309-63	pHS3	AM99-2003	21,24	11
P02-309-111	pHS3	AM99-2003	23,44	22
P02-309-114	pHS3	AM99-2003	22,55	18
P02-311-59	pHS3	AM99-2003	21,83	14
P02-316-25	pHS3	AM99-2003	19,60	2
P02-316-111	pHS3	AM99-2003	24,18	26
P02-318-12	pHS3	AM99-2003	23,27	22
P02-314-4	pHS4	Desiree	19,77	3
P02-313-21	pHS4	AM99-2003	21,08	10
P02-313-42	pHS4	AM99-2003	20,61	8
P02-317-2	pHS4	AM99-2003	20,76	8
P02-317-15	pHS4	AM99-2003	19,45	2

10

Table 7: Analysis of dry matter content in lines over-expressing StGH1 (pHS3) or StGH2 (pHS4) grown in field trial.

15

Analyses of the transgenic lines over expressing the StGH1 or StGH2 genes show an increase in dry matter content compared to its respective parental line. As can be seen in table 7 the dry matter is increased up to 26% in lines over-expressing the StGH1

gene. The increase in dry matter is more pronounced when AM99-2003 is used as parental line. This is due to the fact that AM99-2003 is containing a significant amount of available sugars as a consequence of the high amylose trait (see figure 2). Also in lines over-expressing StGH2 the dry matter is increased. The increase is also in his case higher when AM99-2003 is used as parental line. Lines over-expressing StGH2 have an increase in dry matter of up to 10%.

Example 11

Field-trial of transgenic potato lines

10

Transgenic lines as described in examples 5 to 10 are tested in field trials for the determination of agronomic performance in relation to the parental line and other varieties used for starch production. Starch content, which is a main agronomic factor of importance for crops used for starch processing, can be measured by several different methods.

15

Under water weighing of tubers is performed on a scale in a tub of water. Starch content was determined according to standard procedure. 5 kg potato is used for the measurement and starch content is calculated according to the formula: Starch content in % = (density of potato - 1.01506) / 0.0046051. An increase in starch content is associated with an increase in the density of the sample. An increase of starch in the tubers is associated with an increased dry matter content, which can be measured by comparing the tissue fresh weight to tissue dry weight after extensive water elimination in an oven at 105°C for 16 hours.

25

Starch content can also be measured by enzymatic methods as described under starch content analysis in example 9 and 12.

Example 12

Results in field-trial

30

Five lines over-expressing StGH1 or StGH2 were grown in the field as cuttings. The growth period was within June to September. After harvest the lines were analyzed for dry matter content, starch content and sugar content, see results in tables 8-10.

35

The dry matter content was analyzed by drying 15 g of mashed potatoes (produced in a blender) in a fanned heating oven for 16-18 hours at 105°C. The samples were cooled down to room temperature in an exicator before measurement.

The starch content from field grown tubers was analyzed according to an enzymatic method described in P. Åman et al. Methods in Carbohydrate Chemistry Vol. X. 1994, pp.111-115 by using a thermostable α -amylase. Duplicate analysis was made on ground and dried samples of tubers diluted in ethanol (80%) and digested by thermostable α -amylase and amyloglucosidase. The amount of starch was determined by a glucose oxidase reaction.

The concentration of fructose, glucose and sucrose was determined using gas-liquid chromatography by methods described by Georg Fuchs et al., Swedish J. Agric. Res. 4:49-52, 1974, Quantitative determination of low-molecular carbohydrates in foods by gas-liquid chromatography.

Line No.	Construct	parental line	Dry matter %	Increase in dry matter %
AM99-2003			20,7	
Desiree			21,8	
P02-307-33	pHS3	Desiree	22,6	4
P02-307-80	pHS3	Desiree	22,3	2
P02-309-63	pHS3	AM99-2003	21,9	6
P02-309-106	pHS3	AM99-2003	21,2	2
P02-313-21	pHS4	AM99-2003	22,2	7

Table 8: Analysis of dry matter content in lines over-expressing StGH1 or StGH2 grown in field-trials. Results presented are mean values of two analyses.

As can be seen in table 8 the field grown transgenic lines over-expressing StGH1 or StGH2 show an increase in dry matter of up to 7%. The lines also show an increase in starch content as can be seen in table 9. The highest increase can be seen for lines with AM99-2003 as parental line. This is due to the access of sugars available in the high amylose parental line (figure 2).

Line No.	Construct	parental line	Starch content %	Increase of starch content %
AM99-2003			13,3	
Desiree			16,7	
P02-307-33	pHS3	Desiree	17,5	5
P02-309-63	pHS3	AM99-2003	14,6	10
P02-309-106	pHS3	AM99-2003	14,9	12
P02-313-21	pHS4	AM99-2003	15,4	16

Table 9: Analysis of starch content determined by an enzymatic and a gravimetric method. The lines over-expressing StGH1 (pHS3) or StGH2 (pHS4) were grown in field-trials.

5

Furthermore the sugar concentrations were analyzed in lines with AM99-2003 as parental line. AM99-2003 contains a high fraction of available sugars due to the high amylose trait. In the lines over-expressing the StGH1 or the StGH2 gene the concentration of glucose has been reduced to $1/3$ and sucrose has been reduced to $3/4$ of the amount analyzed in the parental line (see table 10). The lower sugar concentrations in the lines over-expressing the StGH1 or StGH2 gene show that more glucose and sucrose has been incorporated in the starch biosynthesis resulting in an increase in starch content in these transgenic lines.

10

Line No.	Construct	parental line	Fructose % of DM	Glucose % of DM	Decrease in glucose content %	Sucrose % of DM	Decrease in sucrose content %
AM99-2003			0,01	1,2		3,25	
P02-309-63	pHS3	AM99-2003	0,02	0,91	-24	2,74	-16
P02-309-106	pHS3	AM99-2003	0,01	0,62	-48	2,56	-21
P02-313-21	pHS4	AM99-2003	0,01	0,39	-68	2,39	-26

15

Table 10: Analysis of fructose, sucrose and glucose content in lines over-expressing StGH1 (pHS3) or StGH2 (pHS4) grown in field trial. DM = dry matter.

Example 13

Microscopic investigation of lines over-expressing the StGH1 or the StGH2 gene.

Field grown tubers of transgenic lines over-expressing the StGH1 or the StGH2 gene with AM99-2003 as parental line were investigated for starch granule morphology by staining starch with iodine (Lugol's solution (6.7 g/l KI +3.3 g/l I₂) and glycerol 1:1). A piece of a tuber was crushed and a few drops of iodine solution were added. The starch granule structure was analyzed under the microscope.

As can be seen in the figure 10, the starch granules are collapsed towards the interior of the granule in the high amylose parental variety AM99-2003. In contrast to this, the starch granules from the transgenic lines over-expressing the StGH1 or the StGH2 gene are - see figures 11 and 12 - larger and rounded in shape. This is due to the increased starch incorporation in the granules as a result of the over-expression of the StGH1 or the StGH2 gene.

Example 14

Combined expression of genes related to the starch initiation

StGH1 and StGH2 can be combined in different ways. The genes can be combined on the same T-DNA or be located on separate T-DNAs. The genes can be used for co-transformation or be combined by crossing of transgenic lines.

Example 15

Combined constructs for inhibition of SBE1 and SBE2 and over-expression of StGH1 or StGH2

Constructs were made for production of high amylose lines containing high starch content. The constructs pHASHS2, pHASHS4, pHASHS5 and pHASHS6 were made for over-expression of StGH1 or StGH2. All pHASHS constructs also contain fragments of be1 and be2 for down-regulation of respective genes with the antisense technique or RNA interference technique (RNAi). The down-regulation of the SBE1 and SBE2 genes inhibits the amylopectin biosynthesis and directs the starch biosynthesis towards increased amylose production. All constructs are based on the binary vector pSUNAHASmodb. The RNAi constructs are based on vector pHAS8b and the antisense constructs are based on vector pHAS4b.

Plants transferred with pHASHS2, pHASHS4, pHASHS5 or pHASHS6, yielded high-amylose lines. The starch content in the produced transgenic lines was higher than in high-amylose lines not over-expressing the StGH1 or StGH2 gene. The starch content was in the same range as can be seen for pHS3 and pHS4 lines described above.

Example 16

Combined constructs for inhibition of SBE1 and SBE2 and over-expression of StGH1 and StGH2

5

Constructs were made for production of high amylose lines containing high starch content. pHASHS7 and pHASHS8 were made for over-expression of StGH1 and StGH2 together in one plant. All pHASHS constructs also contain fragments of SBE1 and SBE2 for down-regulation of respective genes with the antisense technique or
10 RNA interference technique (RNAi). The down-regulation of the SBE1 and SBE2 genes inhibits the amylopectin biosynthesis and directs the starch biosynthesis towards increased amylose production. All constructs are based on the binary vector pSUNAHASmodb. The RNAi constructs are based on vector pHAS8b and the antisense constructs are based on vector pHAS4b.

15

Plants transformed with pHASHS7 and pHASHS8 yielded high-amylose lines. The starch content in the produced transgenic lines was higher than in high-amylose lines not over-expressing the StGH1 and StGH2 genes together.

20

Example 17

Vector Constructions

Construction of pSUNAHASmodb

25

A binary vector based on pSUN1 with a mutated AHAS gene as selection marker was constructed. The vector was used for further cloning of trait genes.

30

A 608 bp fragment containing the nos promoter was cut out from pGPTV-kan with HindIII and BglII and was ligated to pUC19 (Invitrogen) cut open with HindIII and BamHI.

35

The nos terminator (275bp) was cut out from pGPTV-kan with SstI and EcoRI and ligated to above between the SstI and EcoRI sites. The AHAS gene (S653N) described by Sathasivan et al (1991) was optimised by elimination of the restriction sites HindIII, EcoRV, BamHI, EcoRI and SstI by using QuikChange Multi Site directed Mutagenesis kit (Stratagene). Additional restriction sites, KpnI and SstI was added at the 3' and 5' ends of the gene. The gene was named AtAHASmod (figure 22, SEQ ID NO 16). AtAHASmod was cut with KpnI and SstI (ca 2019bp) and ligated between the nos promoter and nos terminator at KpnI-SstI. The above complex was cut out from pUC19 with HindIII (blunt) and EcoRI (2900bp) and ligated to pSUN1 at EcoRI and SmaI. The vector was named pSUNAHASmodb, see figure 19.

40

Construction of pHAS8b

An RNAi construct pHAS8b with a be1 and be2 fragment (SEQ ID NO 20) for down-regulation of the branching enzyme genes was constructed in the binary vector pSUNAHASmodb (figure 19) based on pSUN1 with a mutated AHAS gene (SEQ ID NO 16) as selection marker. As spacer a fragment of the be2 promoter was used (SEQ ID NO 18). The vector was used for extended cloning with the StGH1 (SEQ ID NO 1) and StGH2 (SEQ ID NO 3) genes.

- 5 NAHASmodb (figure 19) based on pSUN1 with a mutated AHAS gene (SEQ ID NO 16) as selection marker. As spacer a fragment of the be2 promoter was used (SEQ ID NO 18). The vector was used for extended cloning with the StGH1 (SEQ ID NO 1) and StGH2 (SEQ ID NO 3) genes.
- 10 A 400 bp synthetically produced fragment of be2 (200bp) and be1 (200bp) in pBluescript named RNAi420be2be1 (SEQ ID NO 19) was opened with HindIII (blunt) and Sall (3331 bp). A 262bp fragment of the be2 promoter (SEQ ID NO 18), for use as spacer, was digested with BglII (blunt) and Sall and ligated to the vector and named pMA17.

- 15 Again the 400 bp RNAi420be2be1 fragment was used and ligated in inverted direction to pMA17 opened with XhoI (blunt) and KpnI (3582 bp). The construct was named pMA18. pMA18 was digested with SpeI and KpnI (1120 bp) and the fragment was ligated between a gbss promoter and a nos terminator in pUC19 at XbaI and KpnI
- 20 (3924 bp). The construct was named pMA19b. pMA19b was digested with PvuII and HindIII (2390 bp) and ligated to the binary vector pSUNAHASmodb (figure 19) between SphI (blunt) and HindIII (8932 bp). The construct was named pHAS8b, see figure 20.

Construction of pHAS4b

- 25 A vector pHAS4b with an antisense fragment for down-regulation of SBE1 and SBE2 was constructed in the binary vector pSUNAHASmodb (figure 19) based on pSUN1 with a mutated AHAS gene (SEQ ID NO. 16) as selection marker. The vector was used for extended cloning of the StGH1 and the StGH2 genes.

- 30 The antisense fragment of be1 and be2 together with gbss promoter and nos terminator was cut out of pHAbE12A (figure 9, SEQ ID NO 15) with BsrBI and HindIII (4299bp) and ligated to pSUNAHASmodb (figure 19) digested with SphI (blunt) and HindIII (8932bp). The construct was named pHAS4b, see figure 21.

- 35 Construction of pHASHS5

- The StGH1 gene, the gbss promoter and the nos terminator was digested from pHS3 (figure 6) and cloned into the RNAi construct pHAS8b (figure 20) containing fragment
- 40 of be1 and be2.

The StGH1 gene, gbss promoter and nos terminator was cut out from pHS3 with Dral and EcoRV (3160bp). The fragment was ligated to pHAS8b opened up with EcoRV (11403bp). The construct was named pHASHS5, see figure 15.

5

Construction of pHASHS6

The StGH2 gene, the gbss promoter and the nos terminator was digested from pHS4 (figure 7) and cloned into the RNAi construct pHAS8b (figure 20) containing fragments of be1 and be2.

10

pHS4 was digested with SpeI and EcoRI. Two fragments were collected for further cloning, a 2486bp EcoRI-EcoRI fragment and a 1169bp SpeI-EcoRI fragment. pBluescript was digested with SpeI and EcoRI. The digested pBluescript was ligated with the 1169bp (SpeI-EcoRI) fragment. The construct was named pMA15. pMA15 was digested with EcoRI (4127bp) and ligated to the 2486bp EcoRI-EcoRI fragment from pHS4. The construct was named pMA16. A 3689bp fragment was cut out from pMA16 digested with EcoRV and ligated to pHAS8b opened with EcoRV (11403bp). The construct was named pHASHS6, see figure 16.

15

20

Construction of pHASHS2

The StGH1 gene, the gbss promoter and the nos terminator was digested from pHS3 (figure 6) and cloned into pHAS4b (figure 21) containing an antisense fragment of be1 and be2.

25

pHS3 was digested with Dral and EcoRV (3160bp) and ligated to pHAS4b opened up with EcoRV (13224bp). The construct was named pHASHS2, see figure 13.

30 Construction of pHASHS4

The StGH2 gene, the gbss promoter and the nos terminator was digested from pHS4 (figure 7) and cloned into pHAS4b (figure 21) containing an antisense fragment of be1 and be2.

35

A 3689bp fragment was digested from pMA16 with EcoRV (6613bp) (for pMA16 see construction strategy of pHASHS3). pHAS4b was digested with EcoRV (13224bp) and ligated with the above fragment. The construct was named pHASHS4, see figure 14.

Construction of pHASHS7

pHASHS7 was designed to contain antisense fragments of be1 and be2 for inhibition of respective gene together with the two amylose biosynthesis enhancing genes StGH1 and StGH2.

pMA16 (for pMA16 see construction strategy of pHASHS3) was digested with EcoRV. The resulting 3649bp fragment was ligated to pHASHS2 (figure 5) opened with PstI (blunted). The construct was named pHASHS7, see figure 17.

10

Construction of pHASHS8

pHASHS8 was designed to contain a fragment of be1 and be2 for inhibition of respective gene using RNAi (SEQ ID NO 19 to 22) and linked by a spacer (SEQ ID NO 18 or 23) together with two amylose biosynthesis enhancing genes StGH1 and StGH2.

15

pMA16 (for pMA16 see construction strategy of pHASHS3) was digested with EcoRV and ligated to pHASHS5 (figure 15) opened with PstI and blunted. The construct was named pHASHS8, see figure 18.

20

Example 18

Increased solids and improved processing quality of potatoes

In another aspect the invention may be used to increase the solids content of potato varieties that are used for processed potato products or as table potato varieties. The potato genotypes are transformed with gene constructs as described above for the over-expression of a gene coding for a starch biosynthesis enhancing protein. This starch biosynthesis enhancing protein may be derived from genes described above or other plant genes containing the same functional domains.

30

Over-expression of the StGH1 and/or the StGH2 gene in potato plants results in an increase in solids as can be seen in table 7 and 8.

What is claimed:

1. A method of increasing the production of starch in plants comprising culturing a plant with enhanced expression or activity of at least one starch biosynthesis enhancing protein.
2. The method of claim 1, wherein said starch has a high amylose content.
3. The method as claimed in either claim 1 or 2, wherein production of amylose is increased.
4. The method as claimed in any one of claims 1 to 3, wherein said method comprises over-expression of a starch biosynthesis enhancing protein.
5. The method as claimed in claim 4, wherein said protein comprises the SEQ ID NO: 2 or 4 or a protein derived from this sequence by substitution, insertion or deletion of amino acids and which has at least 50% identity at the amino acid level with SEQ ID NO: 2 or 4.
6. The method as claimed in any of claims 1 to 5, wherein the starch biosynthesis enhancing protein is encoded by a nucleic acid sequence selected from the group consisting of:
 - a) a nucleic acid sequence comprising a nucleotide sequence which is at least 60% identical to the nucleic acid sequence of SEQ ID NO: 1 or 3;
 - b) a nucleic acid sequence comprising a fragment of at least 30 nucleotides of a nucleic acid sequence comprising the nucleotide sequence of SEQ ID NO:1 or 3;
 - c) a nucleic acid sequence which encodes a polypeptide comprising an amino acid sequence at least about 60% identical to the amino acid sequence of SEQ ID NO:2 or 4 and
 - d) a nucleic acid sequence which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO: 2 or 4 or wherein the fragment comprises at least 10 contiguous amino acid residues of the amino acid sequence of SEQ ID NO:2 or 4,

51

7. The method as claimed in any one of claims 1 to 6, wherein the starch biosynthesis enhancing protein is encoded by a nucleic acid sequence comprising the nucleotide sequence set forth as SEQ ID NO:1 or SEQ ID NO:3.
- 5 8. The method as claimed in any one of claims 1 to 7, wherein deficiency or decreased activity is achieved by a method selected from the group consisting of:
- 10 a) knock-out of the gene encoding said protein;
- b) mutagenesis of the gene encoding said protein, wherein said mutation can be induced in the coding, non-coding, or regulatory regions of said gene;
- 15 c) expression of an anti-sense RNA, wherein said anti-sense RNA is complementary to at least part of the RNA encoding said protein;
9. A method of producing amylose type starch by culturing a plant which over-expresses SEQ ID NO:1 or 3 or has increased starch biosynthesis enhancing activity under conditions such that the plant produces an increased amount of amylose type starch.
- 20 10. The method of any of the preceeding claims, wherein said plant belongs to the genus *Solanum*.
- 25 11. The method of claim 10, wherein said plant is *Solanum tuberosum*.
12. A nucleic acid sequence SEQ ID NO:1 encoding a starch biosynthesis enhancing protein.
- 30 13. A nucleic acid sequence SEQ ID NO:3 encoding a starch biosynthesis enhancing protein.
14. An amino acid sequence SEQ ID NO:2 having starch biosynthesis enhancing activity.
- 35 15. An amino acid sequence SEQ ID NO:4 having starch biosynthesis enhancing activity.

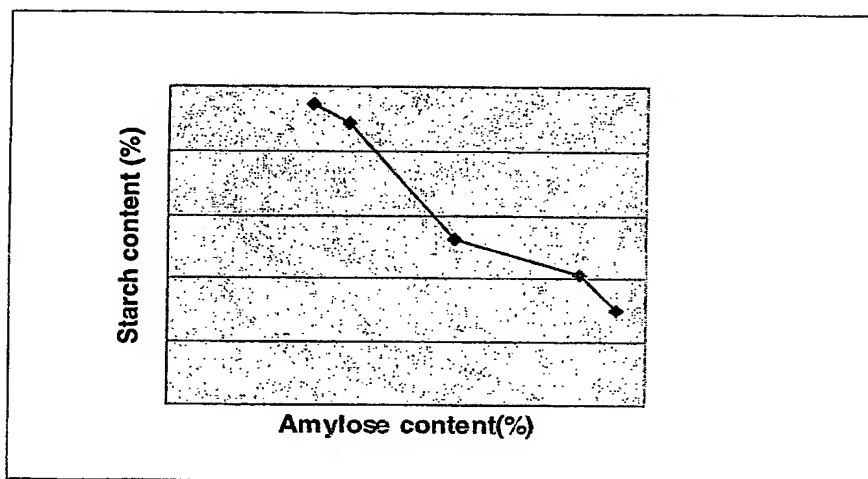
16. A transgenic expression cassette comprising in combination with a regulatory sequence a nucleic acid sequence selected from the group consisting of:
- 5 a) a nucleic acid sequence comprising a nucleotide sequence which is at least 60% identical to the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3,
- b) a nucleic acid sequence comprising a fragment of at least 30 nucleotides of a nucleic acid sequence comprising the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3,
- 10 c) a nucleic acid sequence which encodes a polypeptide comprising an amino acid sequence at least about 60% identical to the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4, or
- d) a nucleic acid sequence which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4
- 15 wherein the fragment comprises at least 10 contiguous amino acid residues of the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4
- wherein said regulatory sequence is capable of mediating expression of said nucleic acid sequence in a plant.
- 20 17. A transgenic expression cassette of claim 16, wherein said regulatory sequence is a promoter sequence heterologous with regard to said nucleic acid sequence.
18. A transgenic expression cassette of claim 16, wherein said regulatory sequence
- 25 is a tuber specific promoter sequence.
19. A transgenic expression cassette of either claim 16, 17 or 18, wherein said nucleic acid sequence is arranged in antisense or sense orientation with regard to said promoter sequence.
- 30 20. A transgenic expression cassette of any of the claims 16 to 19, wherein said nucleic acid sequence encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2 or SEQ ID NO:4.
- 35 21. A transgenic expression cassette of any of the claims 16 to 20, wherein said nucleic acid molecule comprises the nucleotide sequence set forth in SEQ ID NO:1 or SEQ ID NO:3.

53

22. A transgenic expression cassette of any of the claims 16 to 21, wherein said nucleic acid sequence encodes a naturally occurring variant of a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2 or SEQ ID NO:4.
- 5 23. A transgenic host cell transformed with an expression cassette of any of the claims 16 to 22.
24. A transgenic host cell of claim 23, wherein said host cell belongs to the genus *Solanum*.
- 10 25. A transgenic plant comprising an expression cassette of any of claims 16 to 22.
26. A transgenic potato plant comprising an expression cassette of any of claims 16 to 22.
- 15 27. A transgenic potato plant, plant part, seed or tuber comprising an expression cassette of any of claims 16 to 22.

THE PAGE BLANK (MOTO)

Figure 1



JC05 Rec'd PCT/PTO 01 SEP 2005

UNCLASSIFIED (USPTO)

Figure 2

Variety/line	Mother variety	Dry matter %	Fructose % of Dm	Glucose % of Dm	Sucrose % of Dm
Producent		27,8	1,8	2,2	3,1
Prevalent		27,1	1,2	1,7	2,7
Dinamo		27,7	0,8	1,7	2,2
Kuras		27,9	1,8	2,2	2,4
AM98-2012	Producent	22,3	3,1	4,0	4,7
AM98-2019	Prevalent	18,4	2,9	4,4	3,8
AM98-2021	Prevalent	17,1	3,1	5,9	3,5
AM99-2002	Dinamo	19,3	2,4	3,8	3,0
AM99-2003	Dinamo	18,8	2,5	5,5	3,2
AM99-2004	Dinamo	11,7	4,0	6,8	2,1
AM00-2040	Kuras	21,0	5,0	6,3	3,5
AM00-2041	Kuras	19,3	5,6	7,1	3,0

JC05 Rec'd PCT/PTO 31 SEP 2005

THIS PAGE BLANK (COSTO)

Figure 3

3 a) pHS1 for gene-inhibition of StGH1

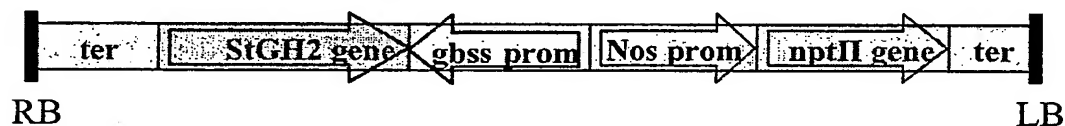
pHS1



5

3 b) pHS2 for gene-inhibition of StGH2

pHS2



3 c) pHS3 for over-expression of StGH1

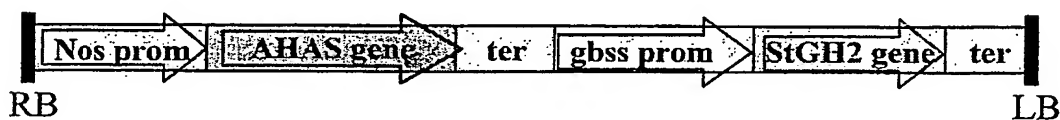
10

pHS3



3 d) pHS4 for over-expression of StGH2

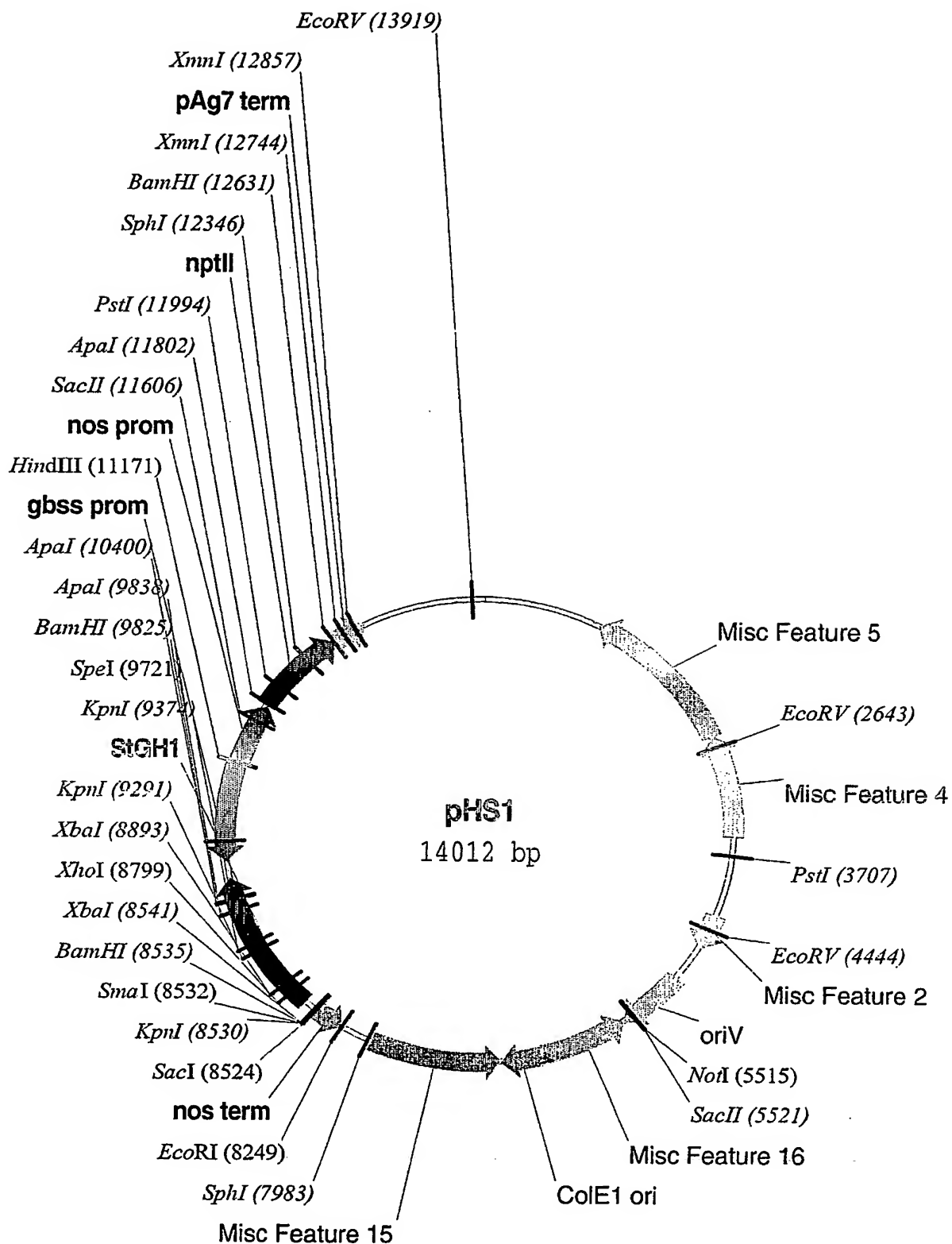
pHS4



JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE BLANK

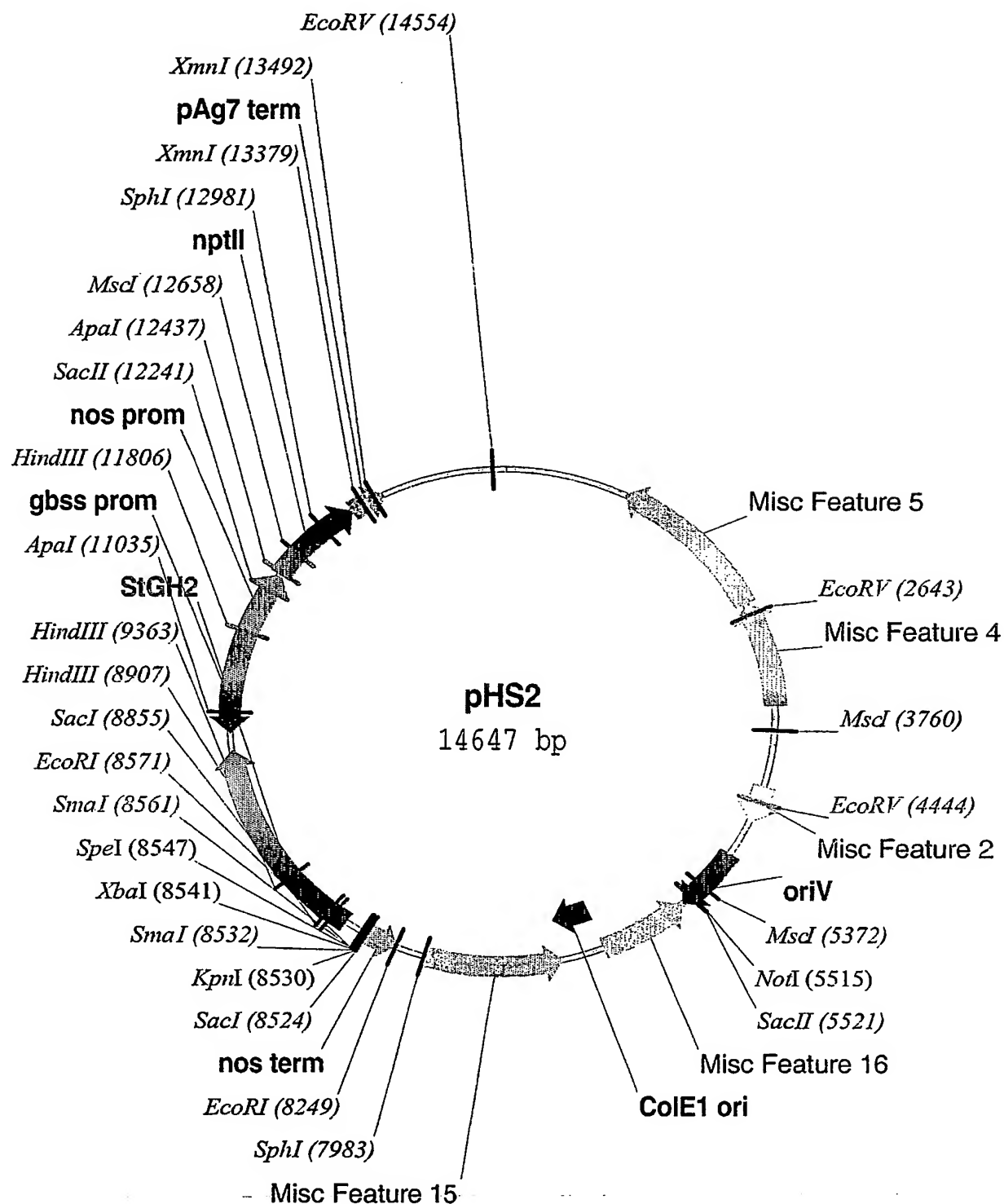
Figure 4



JC05 Rec'd PCT/PTO 01 SEP 2005

USPTO (USPTO)

Figure 5

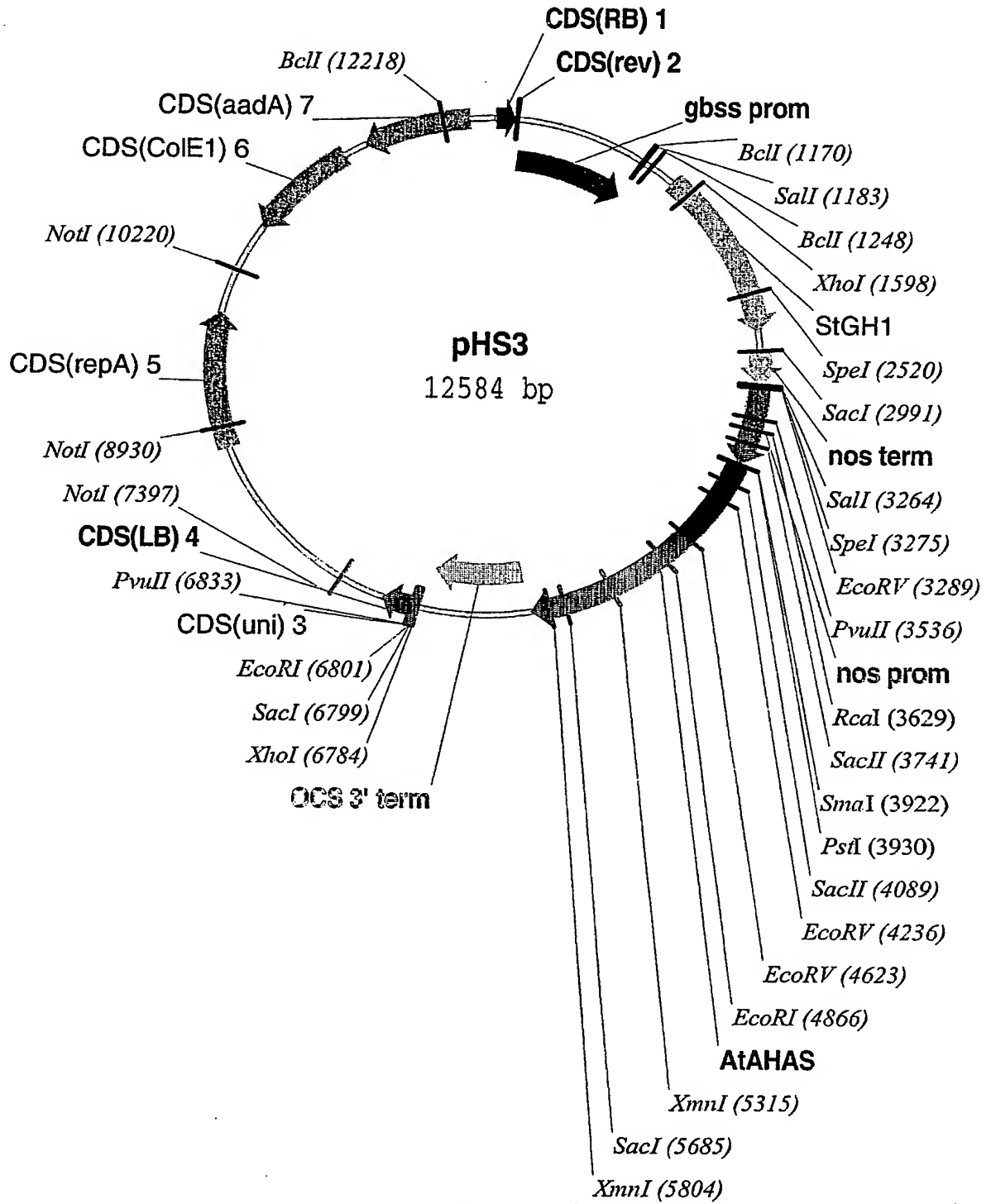


JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE IS BLANK (CONT)

6/23

Figure 6

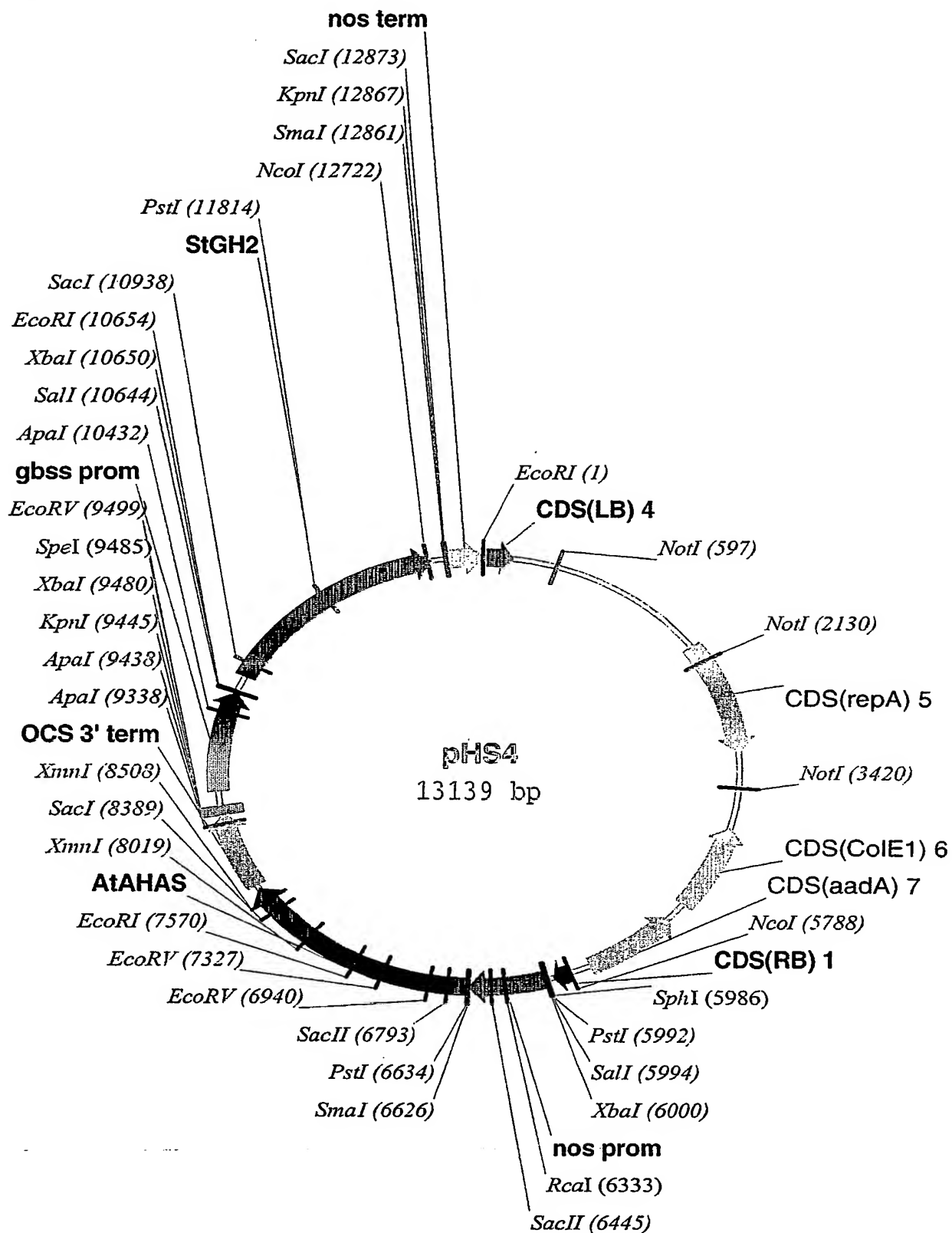


JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE IS BLANK

7/23

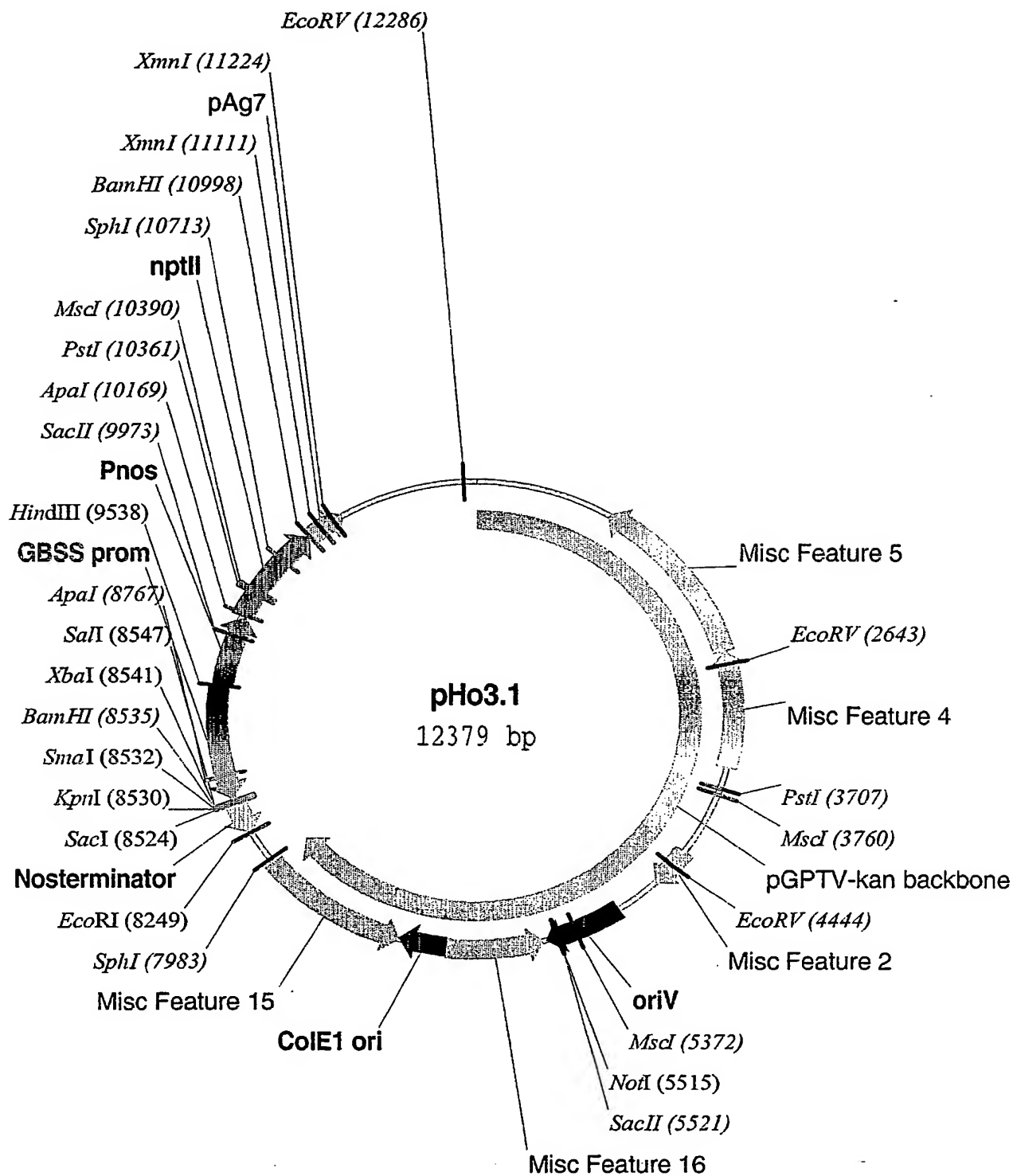
Figure 7



JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE INTENTIONALLY LEFT BLANK

Figure 8

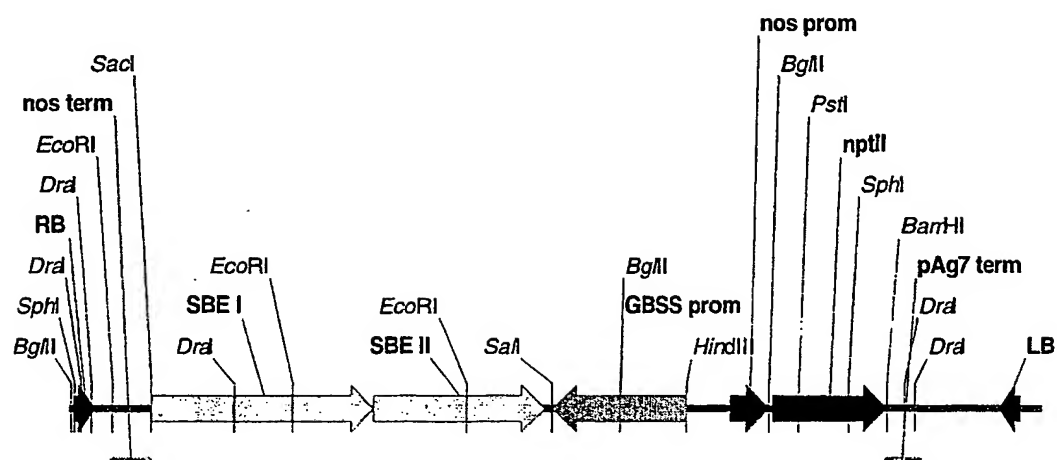


JC05 Rec'd PCT/PTO 01 SEP 2005

AMER DIANK (USPTO)

9/23

Figure 9

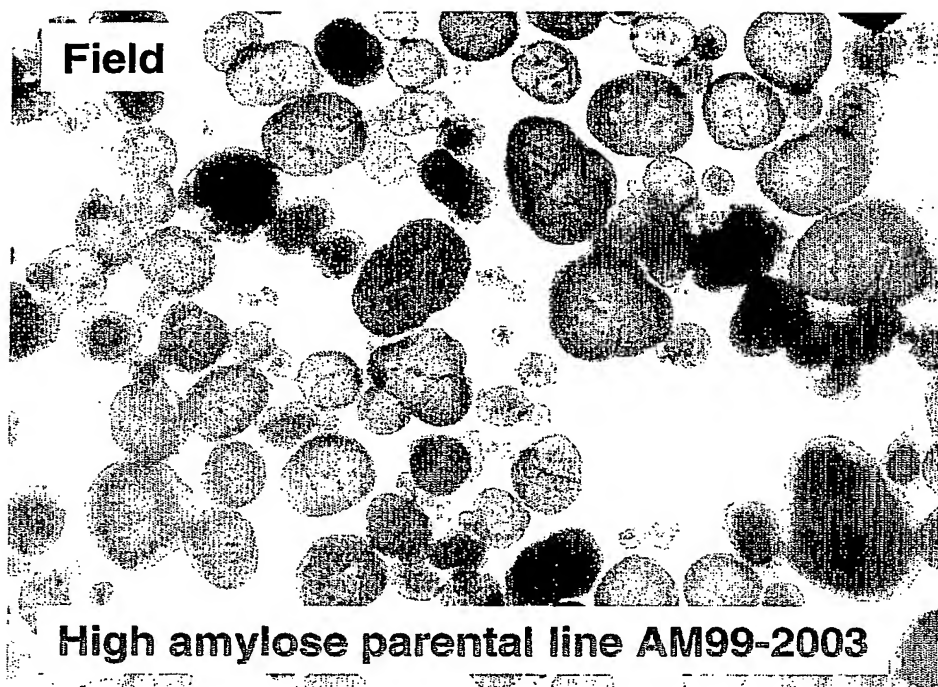


pHAbE12A T-DNA

JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE BLANK (cont.)

Figure 10

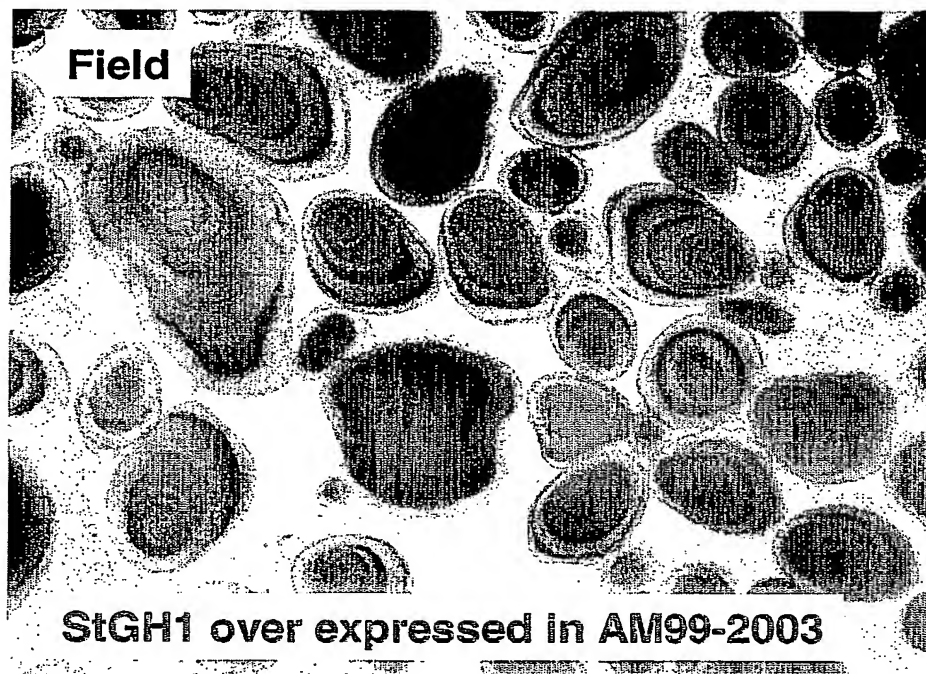


JC05 Rec'd PCT/PTO 01 SEP 2005

END PAGE BLANK (USPTO)

11/23

Figure 11

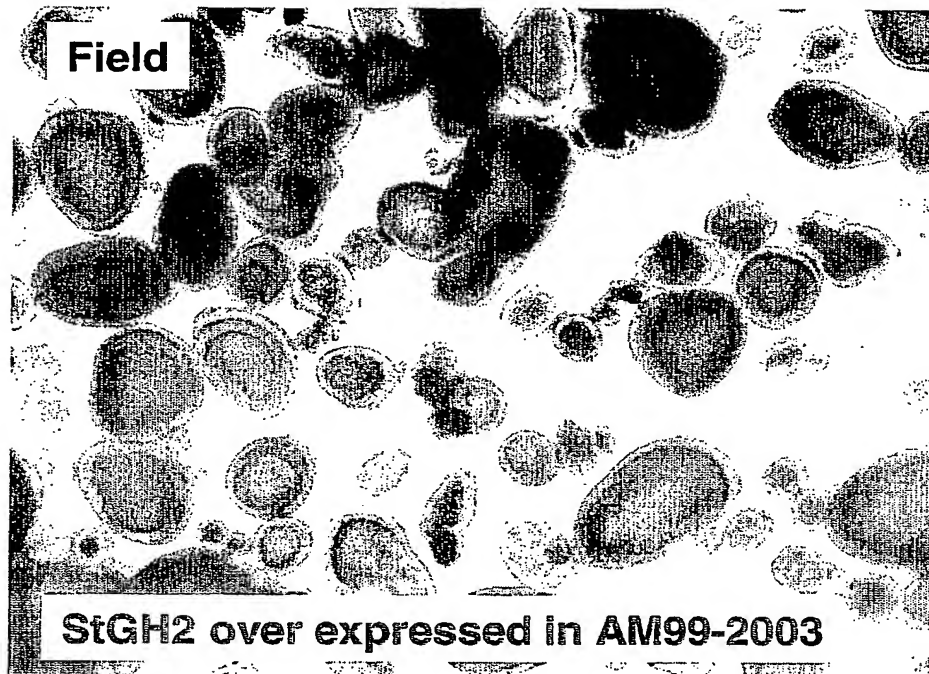


JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE BLANK (CSTO)

12/23

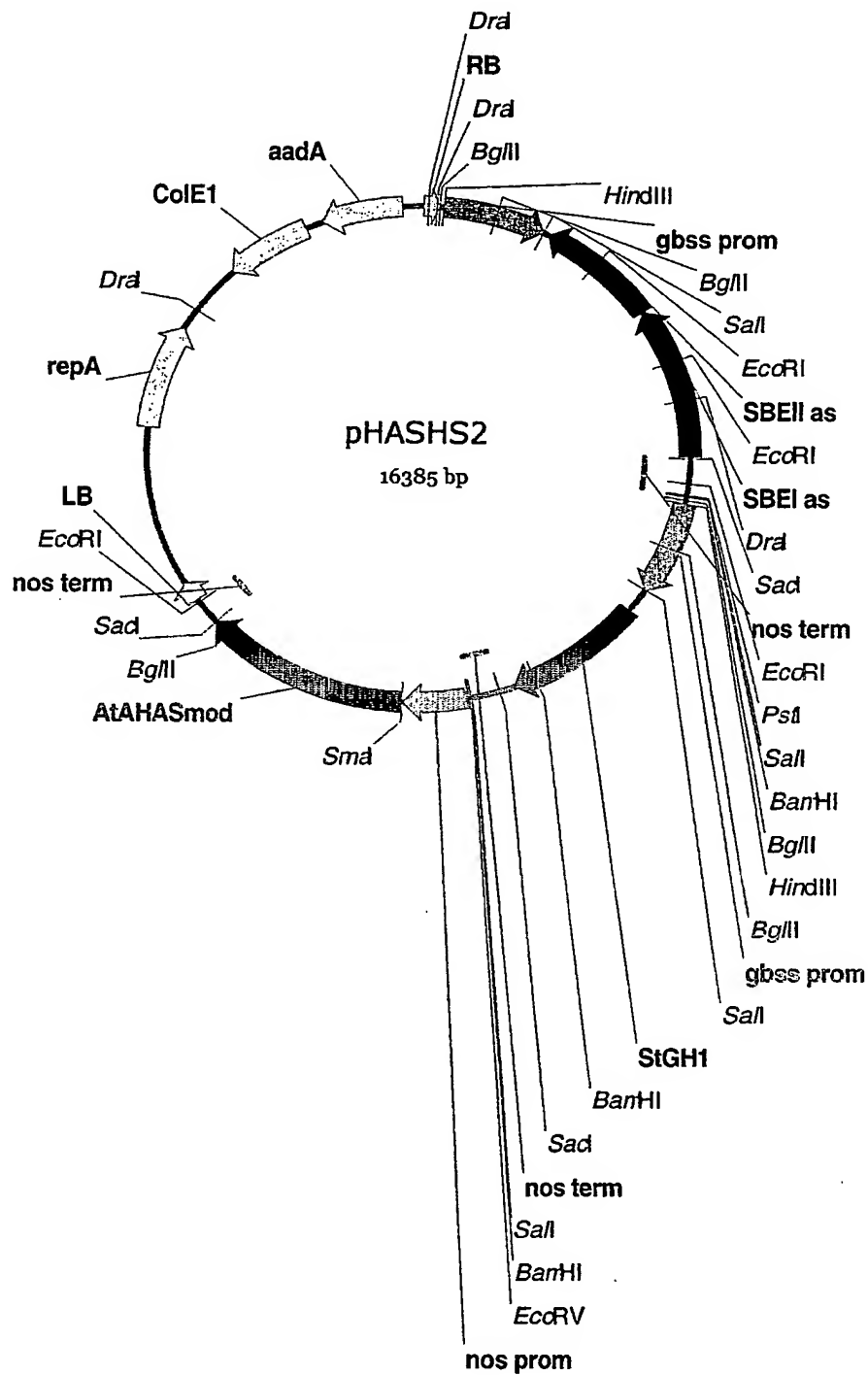
Figure 12



JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE IS BLANK (0000)

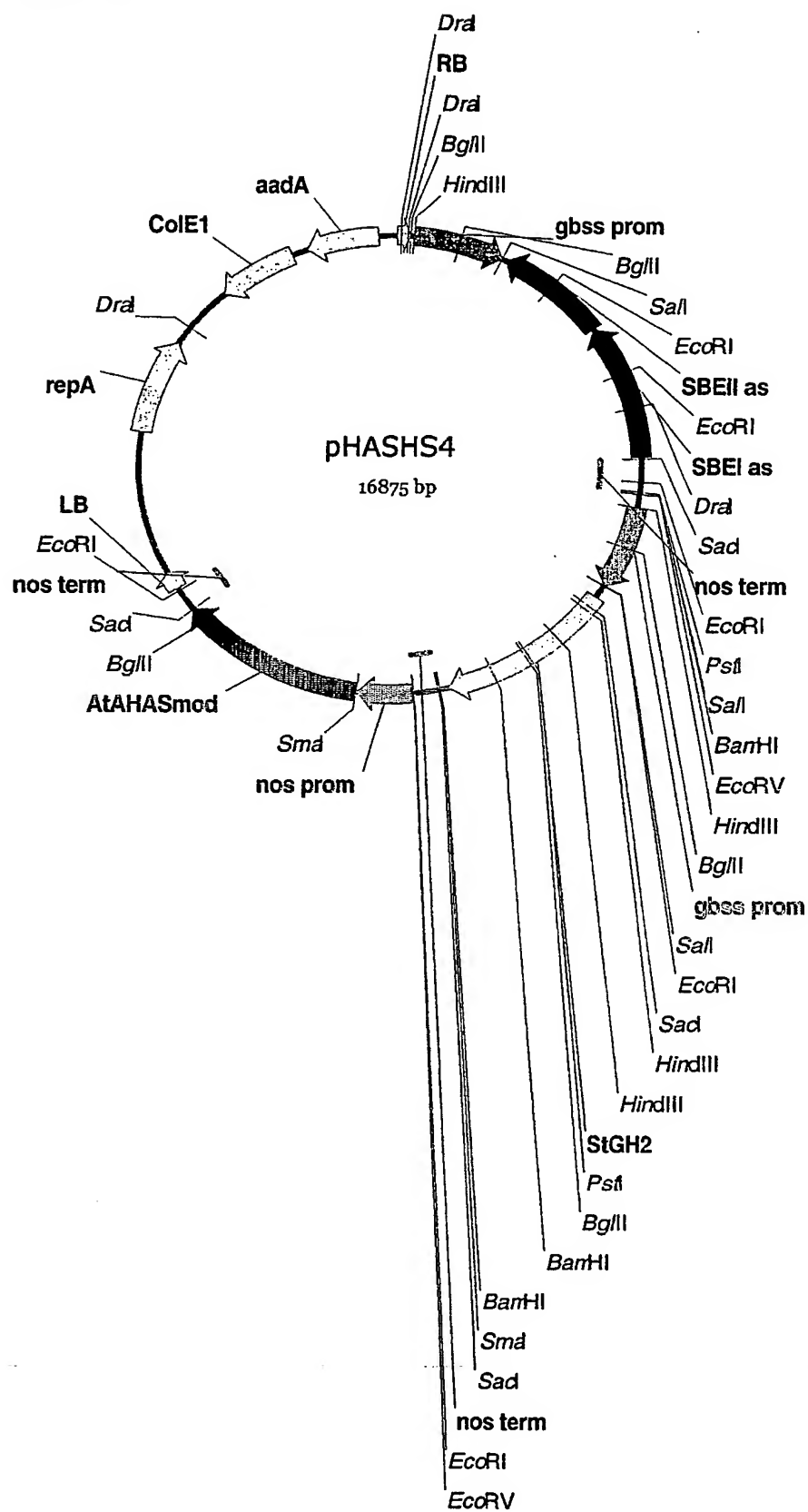
Figure 13



JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE BLANK (CPTO)

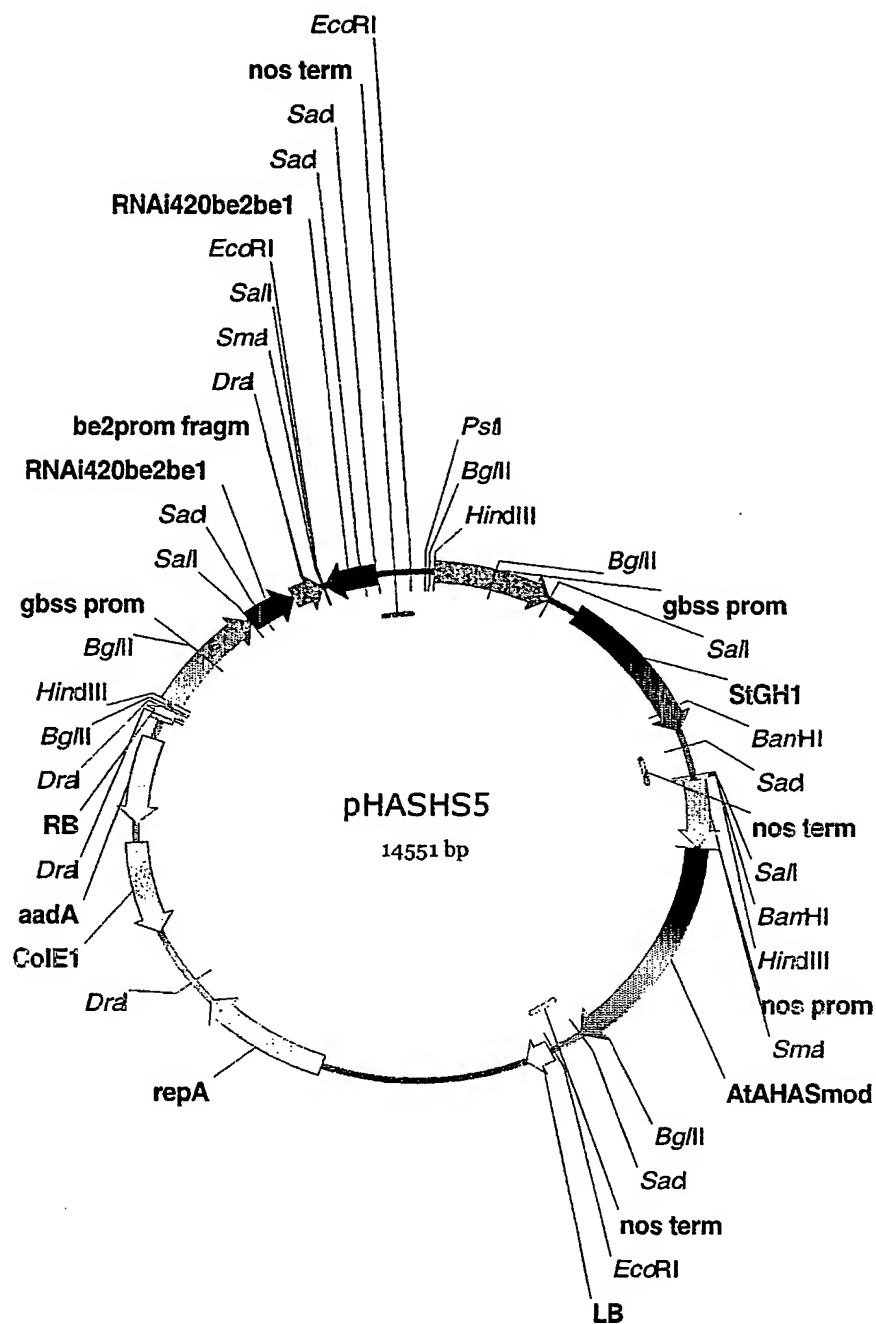
Figure 14



JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE IS BLANK (cont.)

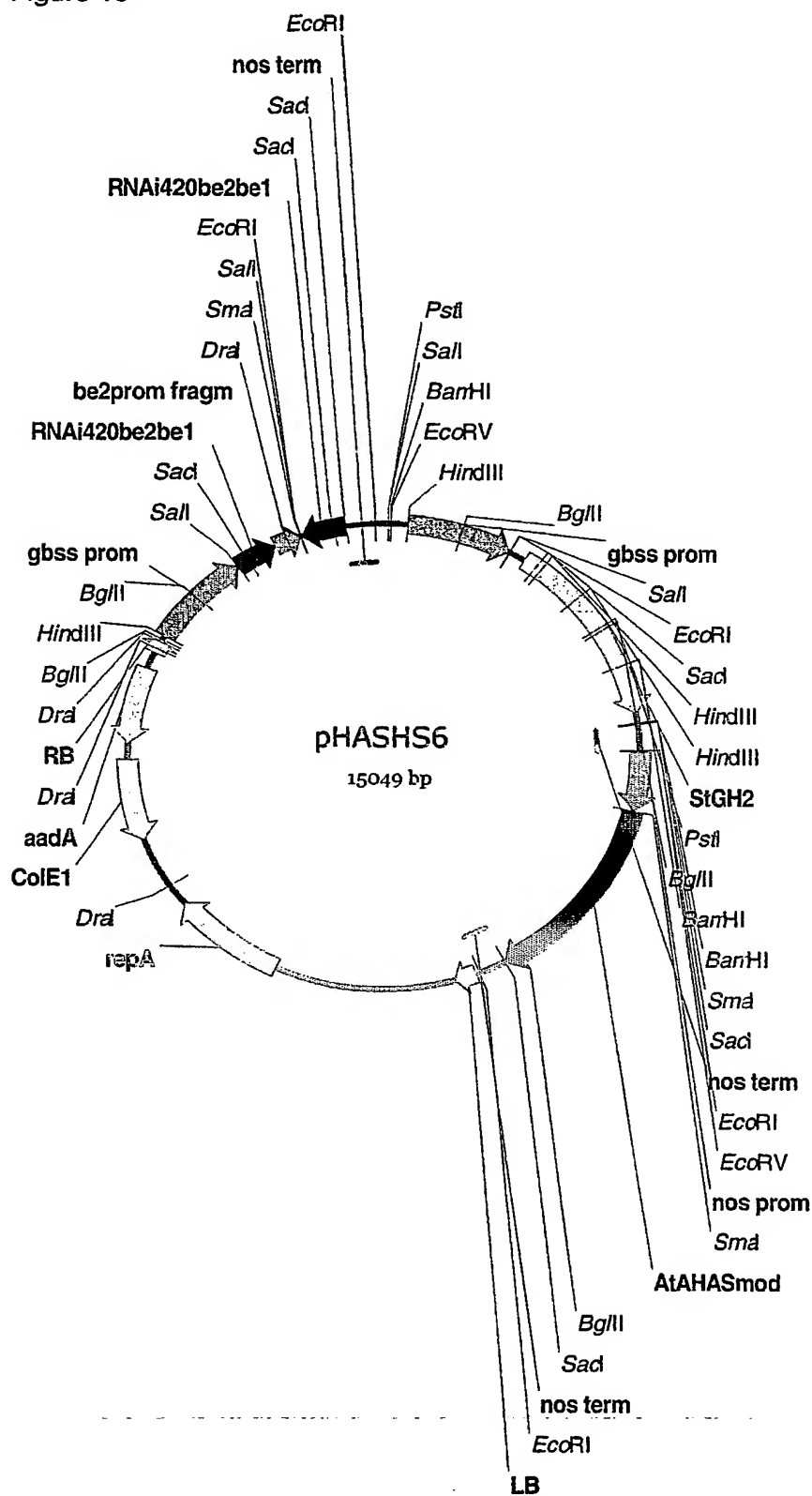
Figure 15



JC05 Rec'd PCT/PTO v1 SEP 2005

THIS PAGE BLANK (cont.)

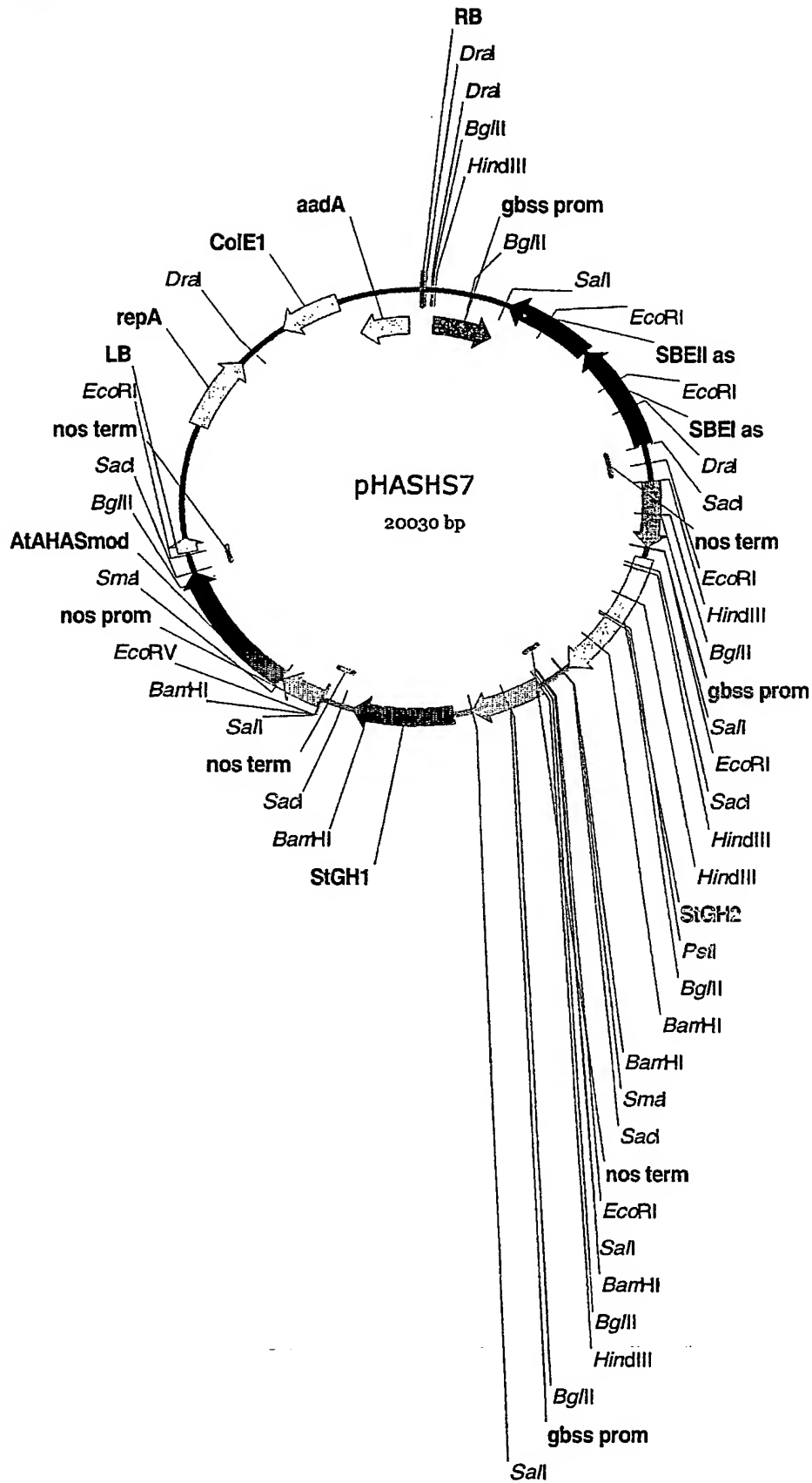
Figure 16



JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE BLANK (USPTO)

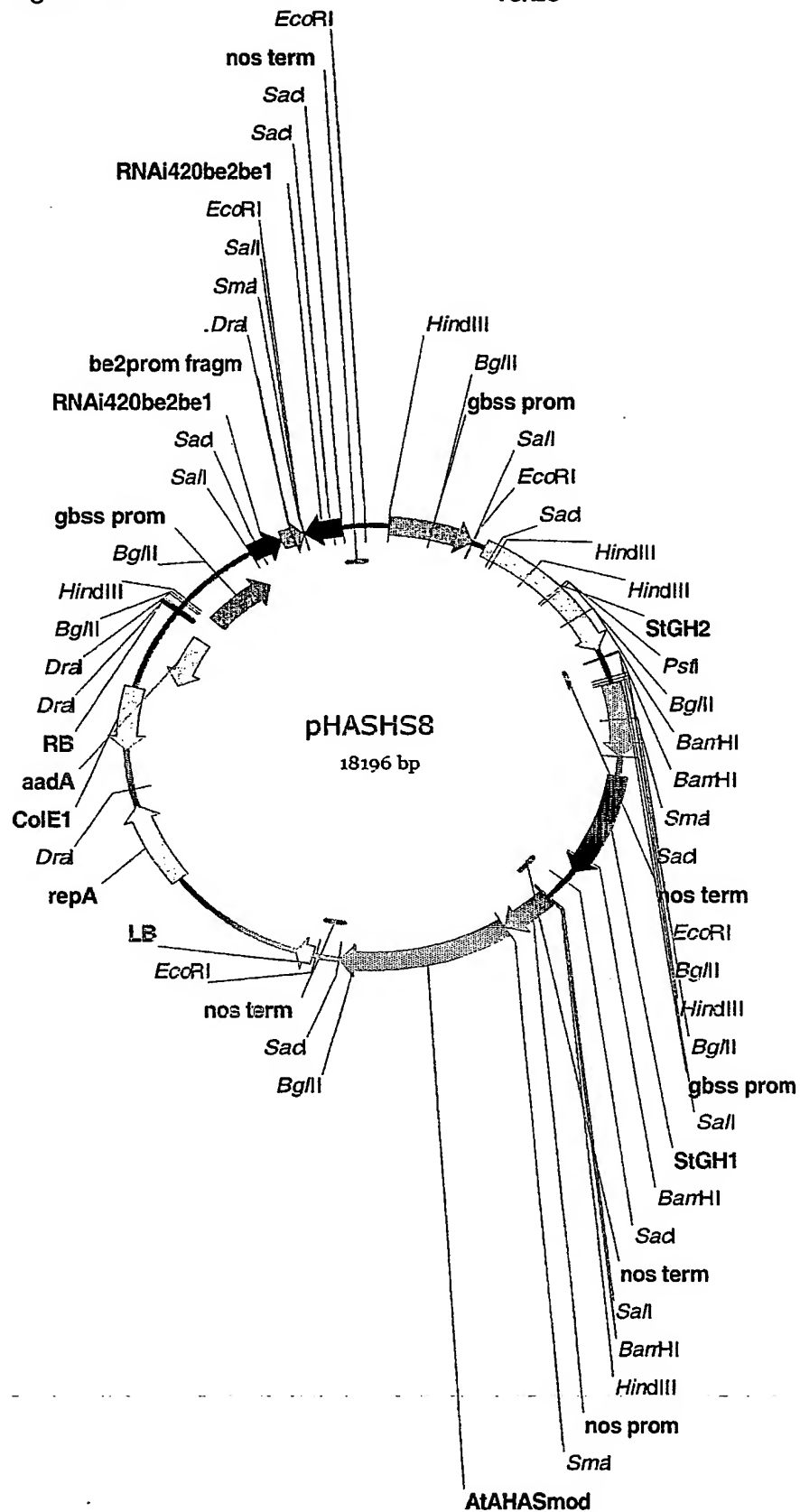
17/23



JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE BLANK (US:10,

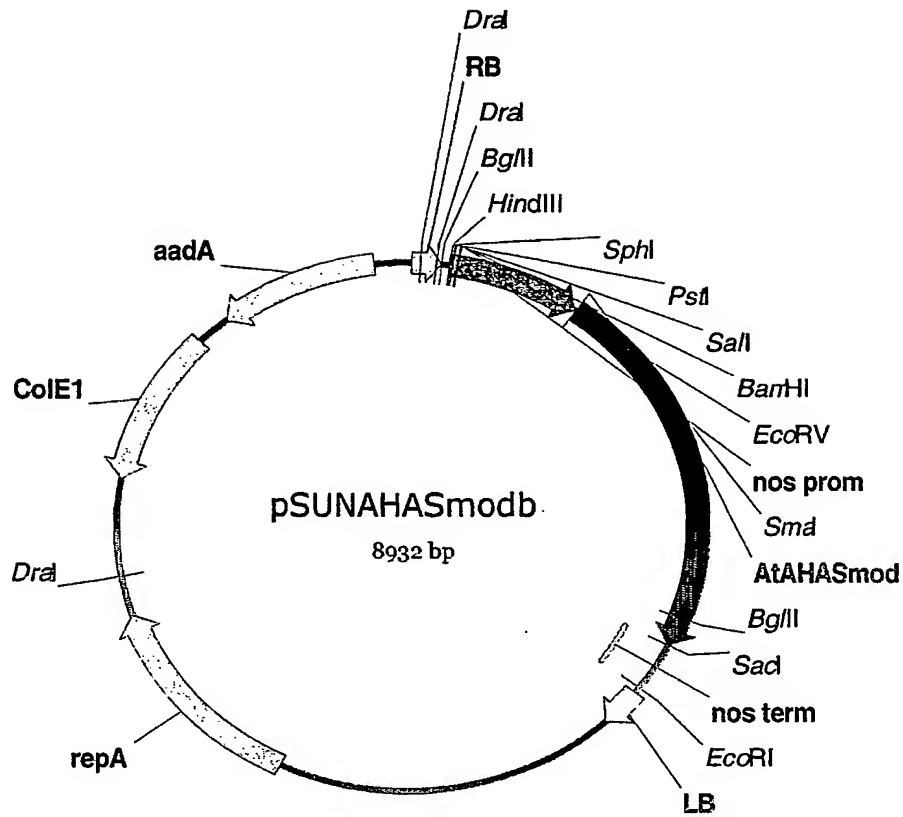
18/23



JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE BLANK (USPTO)

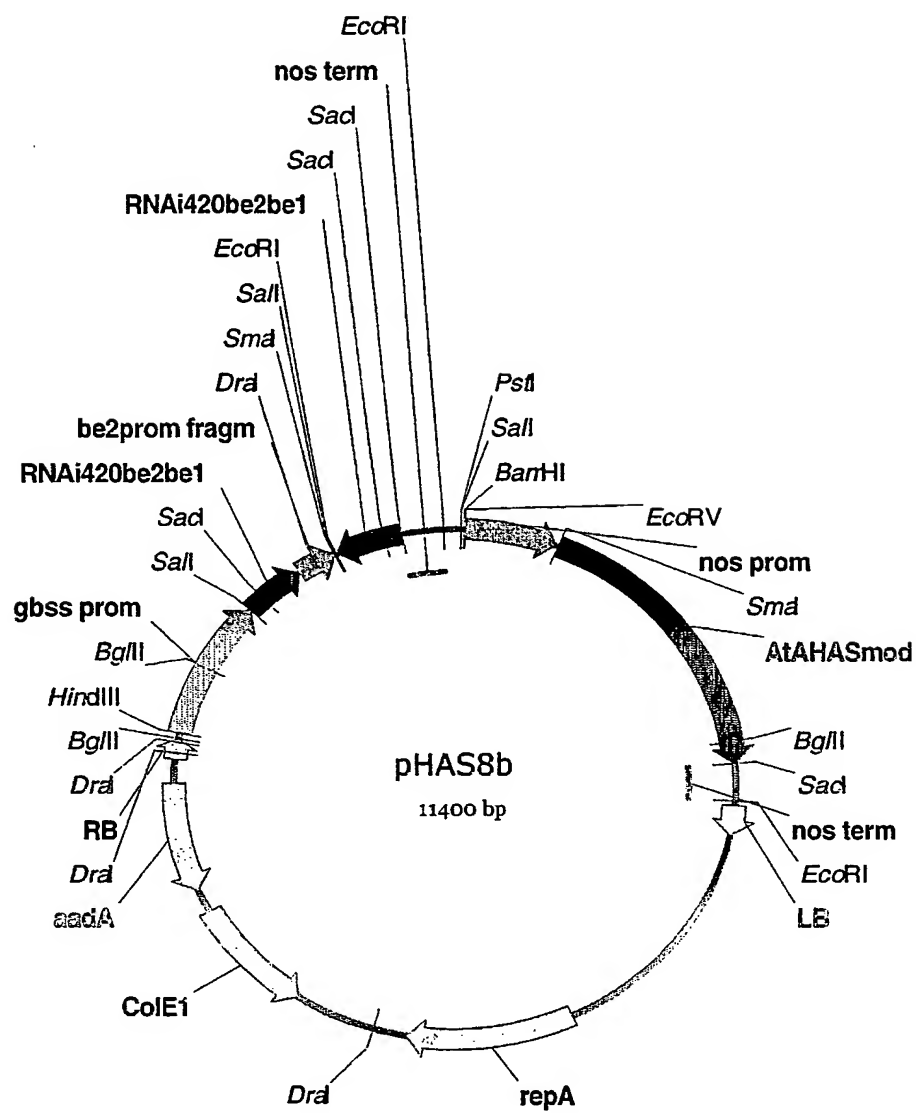
Figure 19



JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE BLANK (cont.)

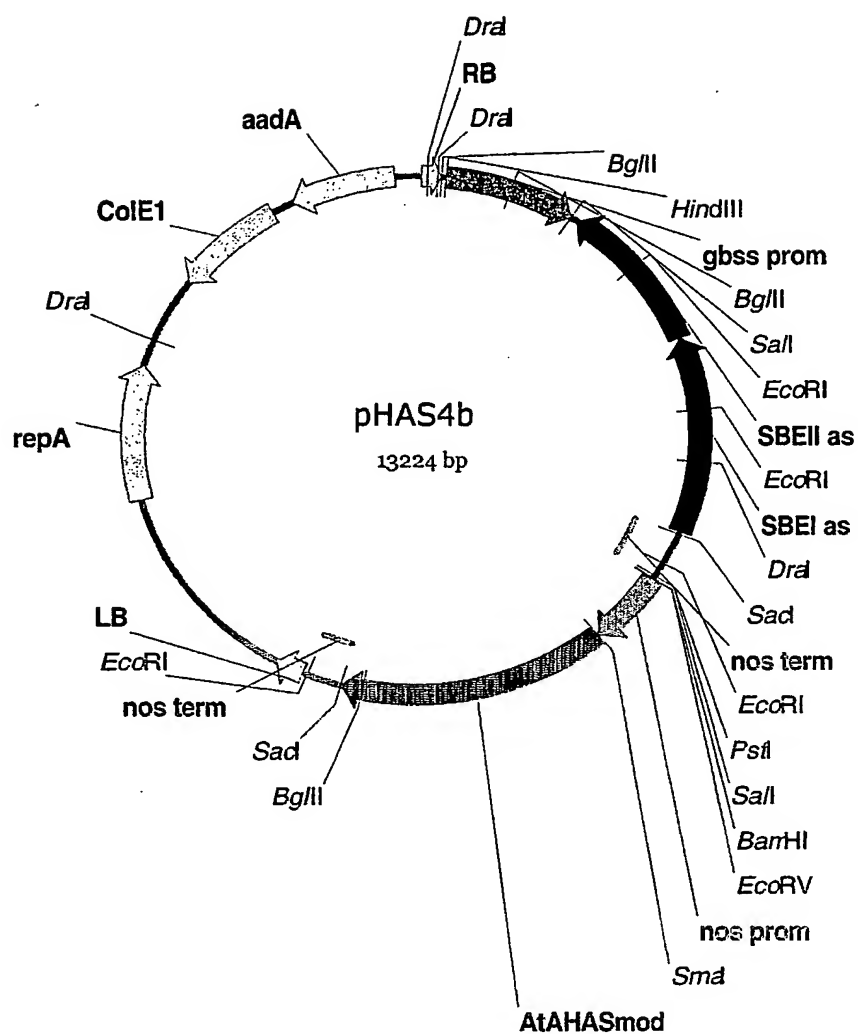
Figure 20



JC05 Rec'd PCT/PTO 01 SEP 2005

INS PAGE BLANK (cont.)

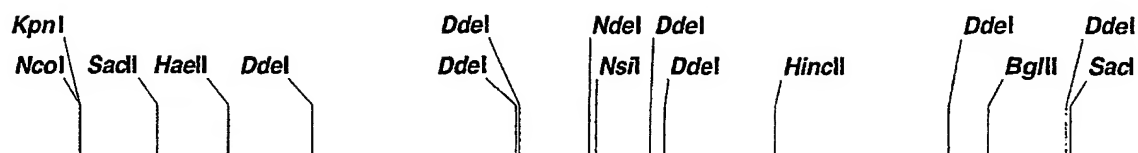
Figure 21



JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE BLANK (US:10,

Figure 22



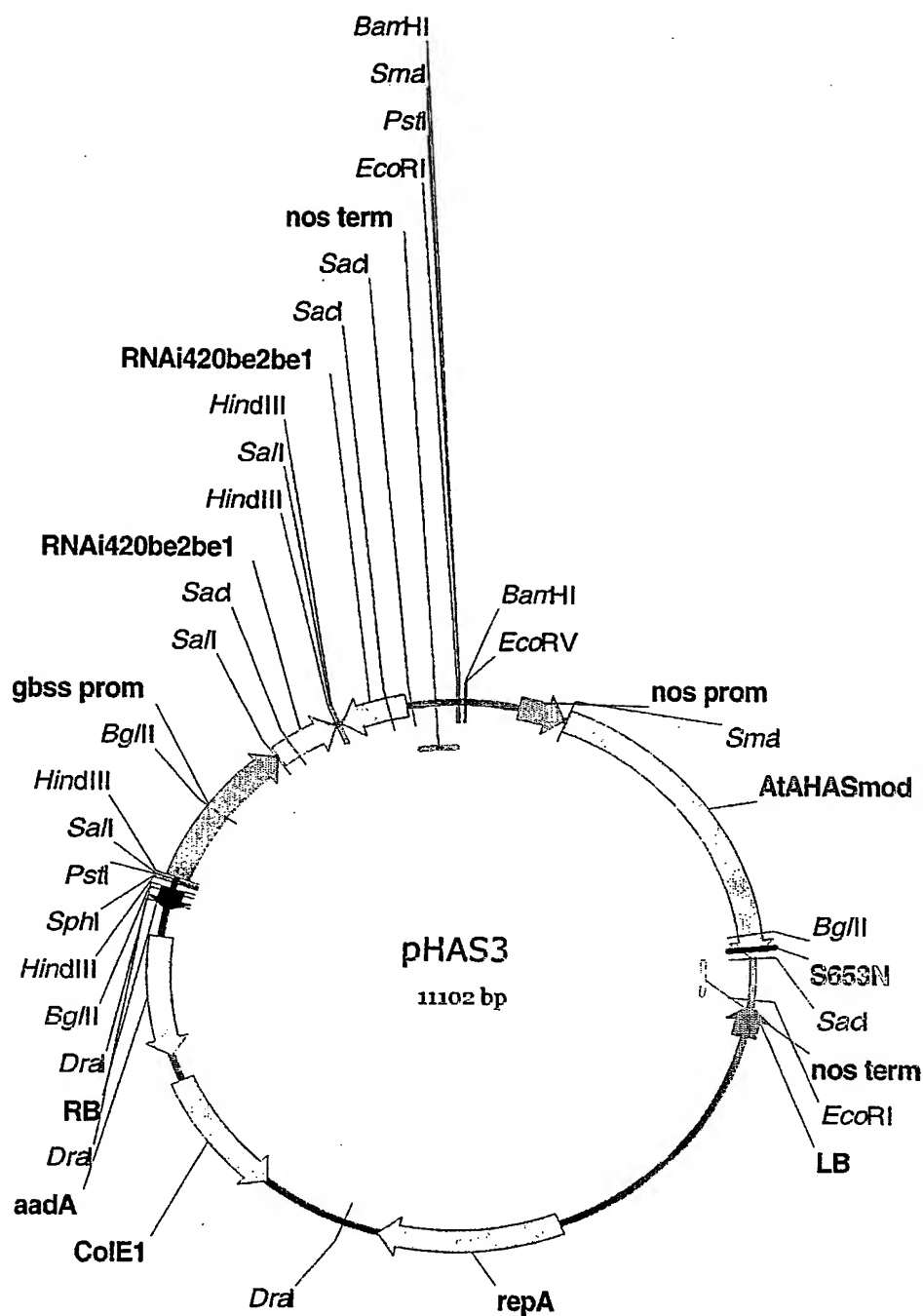
AtAHASmod

2025 bp

JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE BLANK (USPTO)

Figure 23



JC05 Rec'd PCT/PTO 01 SEP 2005

THIS PAGE BLANK (USPTO)

SEQUENCE LISTING

10/547660

<110> BASF Plant Science GmbH

5 <120> Enhanced Amylose Production in Plants

<130> AE884-02

<140> PF00000054331

10 <141> 2003-03-07

<160> 24

<170> PatentIn Ver. 2.1

15 <210> 1

<211> 2084

<212> DNA

<213> Solanum tuberosum

20 <220>

<221> CDS

<222> (302)..(1696)

25 <400> 1

ttttcataaa cttcttcaac tttattccat actcttttat ttatcagctc ctagatcttc 60

ttttttgttt gttgatattc tcttgaaaat tgttcagtga agagttgatc aaagctaaga 120

30 cacaggggct gcggccattt tttcacccga atcttcttct ttattttccg gtgaaagttt 180

gaacgcacac cgttatttct agacagtaga caatgtcaag tgaaaaacat cacaagtttt 240

tgaagatttg taattaatta gttgagattt ttaatttgga ggaaagagaa aaacagagaa 300

35 g atg ata ggg cgg gtg ggc ttg ttg ttg gta ttg ttg ata gca acg acg 349

Met Ile Gly Arg Val Gly Leu Leu Leu Val Leu Leu Ile Ala Thr Thr

1 5 10 15

40 gtg act att ggg gct gaa acg acg acg tta aaa ggg gta aac aga aat 397

Val Thr Ile Gly Ala Glu Thr Thr Thr Leu Lys Gly Val Asn Arg Asn

20 25 30

45 gcg tat gcg act atg atg tat atg gga act ccg aga gac tac gag ttc 445

Ala Tyr Ala Thr Met Met Tyr Met Gly Thr Pro Arg Asp Tyr Glu Phe

35 40 45

50 tac gtg gcg act cga gta atg ctc cga tca ctt acc cgg cta gga gtt 493

Tyr Val Ala Thr Arg Val Met Leu Arg Ser Leu Thr Arg Leu Gly Val

50 55 60

55 gaa gcc gat ctc gtc gtt att gct tca ctt gac gtt cct ctt cgc tgg 541

Glu Ala Asp Leu Val Val Ile Ala Ser Leu Asp Val Pro Leu Arg Trp

65 70 75 80

gtt caa act cta gaa cag gaa gat ggt gct aag gtg gtg aga gtt aaa 589

Val Gln Thr Leu Glu Gln Glu Asp Gly Ala Lys Val Val Arg Val Lys

85 90 95

60 aat ctg aac aat ccg tat tgt atc aac cct aat tgg aga ttc aag ctc 637

Asn Leu Asn Asn Pro Tyr Cys Ile Asn Pro Asn Trp Arg Phe Lys Leu

2

	100	105	110	
5	aca ctg aac aaa ctt tat gcg tgg agc ctc gta aat tat gac agg gtt Thr Leu Asn Lys Leu Tyr Ala Trp Ser Leu Val Asn Tyr Asp Arg Val 115 120 125	685		
10	gtc atg ctt gat gct gac aac ctt ttc ctc cag aaa act gat gaa ctg Val Met Leu Asp Ala Asp Asn Leu Phe Leu Gln Lys Thr Asp Glu Leu 130 135 140	733		
15	ttc caa tgt ggc cag ttt tgt gct gtc ttc att aat ccc tgc atc ttc Phe Gln Cys Gly Gln Phe Cys Ala Val Phe Ile Asn Pro Cys Ile Phe 145 150 155 160	781		
20	cac act ggt ctc ttt gta ttg cag cca tca aaa aag gtg ttc aat gac His Thr Gly Leu Phe Val Leu Gln Pro Ser Lys Lys Val Phe Asn Asp 165 170 175	829		
25	atg atc cat gag ata gag att ggg agg gaa aat caa gac ggt gca gac Met Ile His Glu Ile Glu Ile Gly Arg Glu Asn Gln Asp Gly Ala Asp 180 185 190	877		
30	caa ggt ttt att gga ggc cac ttc cca gat tta ctt gat cgg cca atg Gln Gly Phe Ile Gly Gly His Phe Pro Asp Leu Leu Asp Arg Pro Met 195 200 205	925		
35	ttc cac cct cct ctt aat ggt acc cag ctc cag gga agt tac agg ctt Phe His Pro Pro Leu Asn Gly Thr Gln Leu Gln Gly Ser Tyr Arg Leu 210 215 220	973		
40	cct cta gga tac caa atg gac gcc tct tat tat tat ctc aaa ctc cat Pro Leu Gly Tyr Gln Met Asp Ala Ser Tyr Tyr Tyr Leu Lys Leu His 225 230 235 240	1021		
45	tgg tgg gta cct tgt gga cct aat agt gtc att aca ttt cct ggt gct Trp Ser Val Pro Cys Gly Pro Asn Ser Val Ile Thr Phe Pro Gly Ala 245 250 255	1069		
50	cca tgg tta aaa cca tgg tat tgg tgg tca tgg cct gtc tta ccc ttg Pro Trp Leu Lys Pro Trp Tyr Trp Trp Ser Trp Pro Val Leu Pro Leu 260 265 270	1117		
55	ggc atc cag tgg cat gaa cag cga cgt cta act gtt ggg tat ggt gct Gly Ile Gln Trp His Glu Gln Arg Arg Leu Thr Val Gly Tyr Gly Ala 275 280 285	1165		
60	gag atg ata gca gtg ttg atc caa tct ata ttt tac cta gga ata att Glu Met Ile Ala Val Leu Ile Gln Ser Ile Phe Tyr Leu Gly Ile Ile 290 295 300	1213		
65	gca gtg aca cgc cta gca cgc cca aat tta tca aag ttg tgc tat cgc Ala Val Thr Arg Leu Ala Arg Pro Asn Leu Ser Lys Leu Cys Tyr Arg 305 310 315 320	1261		
70	cat gat gat agc aag agt gcc ttc tta cta cga act ggc ctt aaa ttg His Asp Asp Ser Lys Ser Ala Phe Leu Leu Arg Thr Gly Leu Lys Leu 325 330 335	1309		
75	att gct ata tgg tcc att ctt gct gcc tac aca gtt cct tat ttc gtg Ile Ala Ile Trp Ser Ile Leu Ala Tyr Thr Val Pro Tyr Phe Val 340 345 350	1357		

att cct tgt aca gtt cat cca cta gtt ggc tgg agt ctc tac tta ctc 1405
 Ile Pro Cys Thr Val His Pro Leu Val Gly Trp Ser Leu Tyr Leu Leu
 355 360 365

5

ggc tct ttt tca cta tcc tgt ata aca gtg aat gca ttt ctt ttg ccg 1453
 Gly Ser Phe Ser Leu Ser Cys Ile Thr Val Asn Ala Phe Leu Leu Pro
 370 375 380

10

atg cta cct gtt tta gtc cca tgg att ggg atc ctt ggg gcc ctt ttg 1501
 Met Leu Pro Val Leu Val Pro Trp Ile Gly Ile Leu Gly Ala Leu Leu
 385 390 395 400

15

gtg atg gct tac cct tgg tac aac gac ggt gtt gta aga gca atg gct 1549
 Val Met Ala Tyr Pro Trp Tyr Asn Asp Gly Val Val Arg Ala Met Ala
 405 410 415

20

gta ttt aca tac gcc ttc tgt gct tct cca gca tta tgg atg gca ttg 1597
 Val Phe Thr Tyr Ala Phe Cys Ala Ser Pro Ala Leu Trp Met Ala Leu
 420 425 430

25

gtt aaa atc aag tgt tct ctt cat gtt tca ctt gag agg gaa gga ttc 1645
 Val Lys Ile Lys Cys Ser Leu His Val Ser Leu Glu Arg Glu Gly Phe
 435 440 445

ttg ccc aag ata agt gaa tct aca gca cct gct ggt tct aac aaa ctg 1693
 Leu Pro Lys Ile Ser Glu Ser Thr Ala Pro Ala Gly Ser Asn Lys Leu
 450 455 460

30

tat tgaaagttga aaagttaaag gaatcaacag gagaactaat gcttcagaaa 1746
 Tyr
 465

35

catctccaaa cgtttttgctt aggagacttg gagtctgctt gtgctatcct agctagttgc 1806
 ttcagtctgt gctcttaatt agaatggaat tctgtgagtg ggtttagaat tgggaggatg 1866
 ttttgtgttg tacatggact atctctggtc tcttgaatgc tactccagga aaaagattgt 1926

40

ttctcactta attttttctg ttactaaatt gtatgtggaa taggttcttt aaaatttatt 1986
 catggattta tgttatgtat gctaacagtg taaatattaa gtcctggtga aataagtaat 2046
 tccttattca tacaaaaaaaa aaaaaaaaaa aaaaaaaaa 2084

45

<210> 2
 <211> 465
 <212> PRT

50

<213> Solanum tuberosum
 <400> 2
 Met Ile Gly Arg Val Gly Leu Leu Leu Val Leu Leu Ile Ala Thr Thr
 1 5 10 15
 Val Thr Ile Gly Ala Glu Thr Thr Thr Leu Lys Gly Val Asn Arg Asn
 20 25 30
 Ala Tyr Ala Thr Met Met Tyr Met Gly Thr Pro Arg Asp Tyr Glu Phe
 35 40 45

55

60

4

	Tyr	Val	Ala	Thr	Arg	Val	Met	Leu	Arg	Ser	Leu	Thr	Arg	Leu	Gly	Val	
	50						55					60					
5	Glu	Ala	Asp	Leu	Val	Val	Ile	Ala	Ser	Leu	Asp	Val	Pro	Leu	Arg	Trp	
	65					70					75					80	
	Val	Gln	Thr	Leu	Glu	Gln	Glu	Asp	Gly	Ala	Lys	Val	Val	Arg	Val	Lys	
					85					90					95		
10	Asn	Leu	Asn	Asn	Pro	Tyr	Cys	Ile	Asn	Pro	Asn	Trp	Arg	Phe	Lys	Leu	
				100					105					110			
	Thr	Leu	Asn	Lys	Leu	Tyr	Ala	Trp	Ser	Leu	Val	Asn	Tyr	Asp	Arg	Val	
			115				120					125					
15	Val	Met	Leu	Asp	Ala	Asp	Asn	Leu	Phe	Leu	Gln	Lys	Thr	Asp	Glu	Leu	
	130						135					140					
20	Phe	Gln	Cys	Gly	Gln	Phe	Cys	Ala	Val	Phe	Ile	Asn	Pro	Cys	Ile	Phe	
	145					150				155						160	
	His	Thr	Gly	Leu	Phe	Val	Leu	Gln	Pro	Ser	Lys	Lys	Val	Phe	Asn	Asp	
					165				170					175			
25	Met	Ile	His	Glu	Ile	Glu	Ile	Gly	Arg	Glu	Asn	Gln	Asp	Gly	Ala	Asp	
				180				185					190				
	Gln	Gly	Phe	Ile	Gly	Gly	His	Phe	Pro	Asp	Leu	Leu	Asp	Arg	Pro	Met	
		195					200					205					
30	Phe	His	Pro	Pro	Leu	Asn	Gly	Thr	Gln	Leu	Gln	Gly	Ser	Tyr	Arg	Leu	
	210					215				220							
35	Pro	Leu	Gly	Tyr	Gln	Met	Asp	Ala	Ser	Tyr	Tyr	Tyr	Leu	Lys	Leu	His	
	225				230				235							240	
	Trp	Ser	Val	Pro	Cys	Gly	Pro	Asn	Ser	Val	Ile	Thr	Phe	Pro	Gly	Ala	
					245				250					255			
40	Pro	Trp	Leu	Lys	Pro	Trp	Tyr	Trp	Trp	Ser	Trp	Pro	Val	Leu	Pro	Leu	
				260				265						270			
	Gly	Ile	Gln	Trp	His	Glu	Gln	Arg	Arg	Leu	Thr	Val	Gly	Tyr	Gly	Ala	
		275				280							285				
45	Glu	Met	Ile	Ala	Val	Leu	Ile	Gln	Ser	Ile	Phe	Tyr	Leu	Gly	Ile	Ile	
	290					295			300								
50	Ala	Val	Thr	Arg	Leu	Ala	Arg	Pro	Asn	Leu	Ser	Lys	Leu	Cys	Tyr	Arg	
	305				310				315							320	
	His	Asp	Asp	Ser	Lys	Ser	Ala	Phe	Leu	Leu	Arg	Thr	Gly	Leu	Lys	Leu	
					325				330					335			
55	Ile	Ala	Ile	Trp	Ser	Ile	Leu	Ala	Ala	Tyr	Thr	Val	Pro	Tyr	Phe	Val	
				340			345						350				
	Ile	Pro	Cys	Thr	Val	His	Pro	Leu	Val	Gly	Trp	Ser	Leu	Tyr	Leu	Leu	
		355			360			365									
60	Gly	Ser	Phe	Ser	Leu	Ser	Cys	Ile	Thr	Val	Asn	Ala	Phe	Leu	Leu	Pro	

5

	370	375	380	
	Met Leu Pro Val Leu Val Pro Trp Ile Gly Ile Leu Gly Ala Leu Leu			
	385	390	395	400
5	Val Met Ala Tyr Pro Trp Tyr Asn Asp Gly Val Val Arg Ala Met Ala			
	405	410		415
10	Val Phe Thr Tyr Ala Phe Cys Ala Ser Pro Ala Leu Trp Met Ala Leu			
	420	425		430
	Val Lys Ile Lys Cys Ser Leu His Val Ser Leu Glu Arg Glu Gly Phe			
	435	440		445
15	Leu Pro Lys Ile Ser Glu Ser Thr Ala Pro Ala Gly Ser Asn Lys Leu			
	450	455		460
	Tyr			
20	465			
	<210> 3			
	<211> 2230			
25	<212> DNA			
	<213> Solanum tuberosum			
	<220>			
	<221> CDS			
30	<222> (143)..(2086)			
	<400> 3			
	tatccccaga gaatcagctg aatcaagaac tgatttttag attatgtttt cttgattctt 60			
35	tgaaatggga acttgatttt cagtttttca actcagatgt tgtgttcctt tagctggaaa 120			
	acttgaaaaa ggaaagccca ga atg aga gga agt tta gct ggt gga cca cct 172			
	Met Arg Gly Ser Leu Ala Gly Gly Pro Pro			
	1 5 10			
40	agt cct att gaa cct aga cag agg ctt tct gta ttc act gag gaa aca 220			
	Ser Pro Ile Glu Pro Arg Gln Arg Leu Ser Val Phe Thr Glu Glu Thr			
	15 20 25			
45	agc aaa aga agg ttc ttg aga agt aaa gtt ttc aga gat ggg gag aga 268			
	Ser Lys Arg Arg Phe Leu Arg Ser Lys Val Phe Arg Asp Gly Glu Arg			
	30 35 40			
50	gct ctt cat agt ccc acc aaa aac agg aat ttt acc tgc aag ttc cca 316			
	Ala Leu His Ser Pro Thr Lys Asn Arg Asn Phe Thr Cys Lys Phe Pro			
	45 50 55			
	act gtg aag ctt ata ttg ggt gtt att gct ctg gtt gca att tgg tca 364			
55	Thr Val Lys Leu Ile Leu Gly Val Ile Ala Leu Val Ala Ile Trp Ser			
	60 65 70			
	ctc tgg cat tct cca gca att tat aac acg gaa tac ata tct agt tca 412			
	Leu Trp His Ser Pro Ala Ile Tyr Asn Thr Glu Tyr Ile Ser Ser Ser			
	75 80 85 90			
60	ggc tct cgg gct gct ttg atg cac aga gag tta agt ggt cat tct tca 460			

6

	Gly	Ser	Arg	Ala	Ala	Leu	Met	His	Arg	Glu	Leu	Ser	Gly	His	Ser	Ser	
					95					100					105		
5	gct	gat	caa	cgt	tat	aca	tca	ctt	tta	gat	att	gac	tgg	gac	caa	att	508
	Ala	Asp	Gln	Arg	Tyr	Thr	Ser	Leu	Leu	Asp	Ile	Asp	Trp	Asp	Gln	Ile	
				110					115					120			
10	tcc	caa	gtt	att	gag	aaa	ctg	gcc	gat	agg	cat	gag	tat	cag	ggc	gta	556
	Ser	Gln	Val	Ile	Glu	Lys	Leu	Ala	Asp	Arg	His	Glu	Tyr	Gln	Gly	Val	
			125					130					135				
15	ggg	ata	tta	aac	ttc	aat	gac	agt	gaa	att	gat	cag	ttg	aag	gag	tta	604
	Gly	Ile	Leu	Asn	Phe	Asn	Asp	Ser	Glu	Ile	Asp	Gln	Leu	Lys	Glu	Leu	
		140					145					150					
20	cta	ccg	gac	gct	gag	cat	gta	atc	ttg	aac	ctg	gat	cac	gtc	ccg	aat	652
	Leu	Pro	Asp	Ala	Glu	His	Val	Ile	Leu	Asn	Leu	Asp	His	Val	Pro	Asn	
	155					160					165					170	
25	aat	ata	aca	tgg	gaa	aca	ata	tat	cct	gaa	tgg	ata	gat	gaa	gaa	gaa	700
	Asn	Ile	Thr	Trp	Glu	Thr	Ile	Tyr	Pro	Glu	Trp	Ile	Asp	Glu	Glu	Glu	
					175					180					185		
30	gaa	ttt	gag	gtc	ccc	act	tgt	cct	tct	ctg	ccc	aaa	att	cag	ttt	ccg	748
	Glu	Phe	Glu	Val	Pro	Thr	Cys	Pro	Ser	Leu	Pro	Lys	Ile	Gln	Phe	Pro	
				190					195					200			
35	ggt	aaa	cca	agg	att	gat	ctc	ata	gtt	gta	aag	ctt	cca	tgc	aag	aag	796
	Gly	Lys	Pro	Arg	Ile	Asp	Leu	Ile	Val	Val	Lys	Leu	Pro	Cys	Lys	Lys	
			205					210					215				
40	tct	aag	gac	tgg	tat	aga	gat	gta	gct	cgt	ttt	cac	ttg	cag	ctg	gca	844
	Ser	Lys	Asp	Trp	Tyr	Arg	Asp	Val	Ala	Arg	Phe	His	Leu	Gln	Leu	Ala	
		220					225					230					
45	gca	gca	agg	ctg	gct	gcc	agt	aat	aaa	ggg	tat	cat	cca	ata	cat	gtg	892
	Ala	Ala	Arg	Leu	Ala	Ala	Ser	Asn	Lys	Gly	Tyr	His	Pro	Ile	His	Val	
	235					240				245						250	
50	ctt	ctg	gtt	act	gag	cat	ttc	cca	acc	ccc	aat	ctg	ttc	acc	tgt	aaa	940
	Leu	Leu	Val	Thr	Glu	His	Phe	Pro	Thr	Pro	Asn	Leu	Phe	Thr	Cys	Lys	
				255						260					265		
55	gag	tta	gtt	gta	cgt	gaa	ggc	aat	gca	tgg	cta	tat	gaa	cct	aat	ctg	988
	Glu	Leu	Val	Val	Arg	Glu	Gly	Asn	Ala	Trp	Leu	Tyr	Glu	Pro	Asn	Leu	
				270				275						280			
60	aac	act	tta	aga	gag	aag	ctc	cac	ctc	cct	gtt	ggg	tca	tgt	gaa	ctt	1036
	Asn	Thr	Leu	Arg	Glu	Lys	Leu	His	Leu	Pro	Val	Gly	Ser	Cys	Glu	Leu	
			285					290					295				
65	gca	gtt	cct	ctc	aag	gct	aaa	gca	aat	tgg	cac	tct	gga	aat	gta	aga	1084
	Ala	Val	Pro	Leu	Lys	Ala	Lys	Ala	Asn	Trp	His	Ser	Gly	Asn	Val	Arg	
		300					305					310					
70	cga	gaa	gcc	tat	gca	act	att	ctc	cac	tca	gca	aat	ttt	tat	gta	tgt	1132
	Arg	Glu	Ala	Tyr	Ala	Thr	Ile	Leu	His	Ser	Ala	Asn	Phe	Tyr	Val	Cys	
	315					320					325					330	
75	gga	gcc	ata	gct	gca	gca	cag	agt	att	cgc	ttg	gca	ggg	tca	acc	cga	1180
	Gly	Ala	Ile	Ala	Ala	Ala	Gln	Ser	Ile	Arg	Leu	Ala	Gly	Ser	Thr	Arg	

	335	340	345	
5	gat ctt gtg ata ctt gtt gat gag act atc agt gac tac cac agg ggt Asp Leu Val Ile Leu Val Asp Glu Thr Ile Ser Asp Tyr His Arg Gly 350 355 360	1228		
10	ggt tta gag gct gcc gga tgg aag atc cac acg ata aag aga ata agg Gly Leu Glu Ala Ala Gly Trp Lys Ile His Thr Ile Lys Arg Ile Arg 365 370 375	1276		
15	aat cct aaa gct gaa cag gat gcc tac aat gag tgg aac tat agc aaa Asn Pro Lys Ala Glu Gln Asp Ala Tyr Asn Glu Trp Asn Tyr Ser Lys 380 385 390	1324		
20	ttt cgt ctc tgg cag ctg aca gat tat gac aaa atc atc ttc att gat Phe Arg Leu Trp Gln Leu Thr Asp Tyr Asp Lys Ile Ile Phe Ile Asp 395 400 405 410	1372		
25	gcg gat ttg ttg ata ctg aga aat att gat ttt ctc ttt gag atg cct Ala Asp Leu Leu Ile Leu Arg Asn Ile Asp Phe Leu Phe Glu Met Pro 415 420 425	1420		
30	gaa ata act gca ata gga aat aat gca acc ctt ttt aat tca ggc gtg Glu Ile Thr Ala Ile Gly Asn Asn Ala Thr Leu Phe Asn Ser Gly Val 430 435 440	1468		
35	atg gtc gtt gaa cca tca aat tgc aca ttt cag ctg ttg atg gat cat Met Val Val Glu Pro Ser Asn Cys Thr Phe Gln Leu Leu Met Asp His 445 450 455	1516		
40	atc aat gag att gaa tca tac aat ggt ggg gat cag ggg tat ttg aac Ile Asn Glu Ile Glu Ser Tyr Asn Gly Gly Asp Gln Gly Tyr Leu Asn 460 465 470	1564		
45	gaa ata ttc acc tgg tgg cat agg atc cca aaa cac atg aac ttt ttg Glu Ile Phe Thr Trp Trp His Arg Ile Pro Lys His Met Asn Phe Leu 475 480 485 490	1612		
50	aaa cat tat tgg gaa ggt gat gag gag gag aag aag caa atg aaa aca Lys His Tyr Trp Glu Gly Asp Glu Glu Glu Lys Lys Gln Met Lys Thr 495 500 505	1660		
55	cgt ctt ttt ggt gct gat cct cca gtt ctc tat gtt ctg cac tac tta Arg Leu Phe Gly Ala Asp Pro Pro Val Leu Tyr Val Leu His Tyr Leu 510 515 520	1708		
60	ggc ctg aaa cct tgg tta tgc ttc agg gac tac gat tgc aac tgg aat Gly Leu Lys Pro Trp Leu Cys Phe Arg Asp Tyr Asp Cys Asn Trp Asn 525 530 535	1756		
65	gtg ggt aag ttg cag gag ttt gca agt gat gtg gca cac agg acg tgg Val Gly Lys Leu Gln Glu Phe Ala Ser Asp Val Ala His Arg Thr Trp 540 545 550	1804		
70	tgg aag gtc cat gat gcc atg cca gat aac tta cat aaa tat tgt ttg Trp Lys Val His Asp Ala Met Pro Asp Asn Leu His Lys Tyr Cys Leu 555 560 565 570	1852		
75	ctt agg tct aaa cag aag gct gca cta gag tgg gat cga aga gaa gct Leu Arg Ser Lys Gln Lys Ala Ala Leu Glu Trp Asp Arg Arg Glu Ala 575 580 585	1900		

gag aaa gct aac ttt tca gat ggt cat tgg aag atc aaa ata aag gac 1948
 Glu Lys Ala Asn Phe Ser Asp Gly His Trp Lys Ile Lys Ile Lys Asp
 590 595 600

5 cca cgt ttg gag act tgt tat gaa gaa ttt tgc ttc tgg gaa agc atg 1996
 Pro Arg Leu Glu Thr Cys Tyr Glu Glu Phe Cys Phe Trp Glu Ser Met
 605 610 615

10 tta tgg cac tgg ggt gaa aca aac tgg aca gat aat gcc acc tct tca 2044
 Leu Trp His Trp Gly Glu Thr Asn Trp Thr Asp Asn Ala Thr Ser Ser
 620 625 630

15 cca aca cct ccc atg gtc aat act gct tca ctt tct tct ttg 2086
 Pro Thr Pro Pro Met Val Asn Thr Ala Ser Leu Ser Ser Leu
 635 640 645

taactaggag gttctgttac tatacctcgc tagtctgtaa gtaatacaga gtcacaactt 2146

20 gaatgtcaac cagtgccatt agtatacatg gtgtcaacac tttccccacc cttgaaaaaa 2206
 aaaaaaaaaa aaaaaaaaaa aaaa 2230

25 <210> 4
 <211> 648
 <212> PRT
 <213> Solanum tuberosum

30 <400> 4
 Met Arg Gly Ser Leu Ala Gly Gly Pro Pro Ser Pro Ile Glu Pro Arg
 1 5 10 15

35 Gln Arg Leu Ser Val Phe Thr Glu Glu Thr Ser Lys Arg Arg Phe Leu
 20 25 30

Arg Ser Lys Val Phe Arg Asp Gly Glu Arg Ala Leu His Ser Pro Thr
 35 40 45

40 Lys Asn Arg Asn Phe Thr Cys Lys Phe Pro Thr Val Lys Leu Ile Leu
 50 55 60

Gly Val Ile Ala Leu Val Ala Ile Trp Ser Leu Trp His Ser Pro Ala
 65 70 75 80

45 Ile Tyr Asn Thr Glu Tyr Ile Ser Ser Ser Gly Ser Arg Ala Ala Leu
 85 90 95

50 Met His Arg Glu Leu Ser Gly His Ser Ser Ala Asp Gln Arg Tyr Thr
 100 105 110

Ser Leu Leu Asp Ile Asp Trp Asp Gln Ile Ser Gln Val Ile Glu Lys
 115 120 125

55 Leu Ala Asp Arg His Glu Tyr Gln Gly Val Gly Ile Leu Asn Phe Asn
 130 135 140

Asp Ser Glu Ile Asp Gln Leu Lys Glu Leu Leu Pro Asp Ala Glu His
 145 150 155 160

60 Val Ile Leu Asn Leu Asp His Val Pro Asn Asn Ile Thr Trp Glu Thr

	165							170					175				
	Ile	Tyr	Pro	Glu	Trp	Ile	Asp	Glu	Glu	Glu	Glu	Phe	Glu	Val	Pro	Thr	
				180					185					190			
5	Cys	Pro	Ser	Leu	Pro	Lys	Ile	Gln	Phe	Pro	Gly	Lys	Pro	Arg	Ile	Asp	
			195					200					205				
10	Leu	Ile	Val	Val	Lys	Leu	Pro	Cys	Lys	Lys	Ser	Lys	Asp	Trp	Tyr	Arg	
		210					215					220					
	Asp	Val	Ala	Arg	Phe	His	Leu	Gln	Leu	Ala	Ala	Ala	Arg	Leu	Ala	Ala	
	225					230					235					240	
15	Ser	Asn	Lys	Gly	Tyr	His	Pro	Ile	His	Val	Leu	Leu	Val	Thr	Glu	His	
					245					250					255		
	Phe	Pro	Thr	Pro	Asn	Leu	Phe	Thr	Cys	Lys	Glu	Leu	Val	Val	Arg	Glu	
20				260					265					270			
	Gly	Asn	Ala	Trp	Leu	Tyr	Glu	Pro	Asn	Leu	Asn	Thr	Leu	Arg	Glu	Lys	
			275					280					285				
25	Leu	His	Leu	Pro	Val	Gly	Ser	Cys	Glu	Leu	Ala	Val	Pro	Leu	Lys	Ala	
		290					295					300					
	Lys	Ala	Asn	Trp	His	Ser	Gly	Asn	Val	Arg	Arg	Glu	Ala	Tyr	Ala	Thr	
	305					310					315					320	
30	Ile	Leu	His	Ser	Ala	Asn	Phe	Tyr	Val	Cys	Gly	Ala	Ile	Ala	Ala	Ala	
					325					330				335			
	Gln	Ser	Ile	Arg	Leu	Ala	Gly	Ser	Thr	Arg	Asp	Leu	Val	Ile	Leu	Val	
35				340					345					350			
	Asp	Glu	Thr	Ile	Ser	Asp	Tyr	His	Arg	Gly	Gly	Leu	Glu	Ala	Ala	Gly	
			355					360					365				
40	Trp	Lys	Ile	His	Thr	Ile	Lys	Arg	Ile	Arg	Asn	Pro	Lys	Ala	Glu	Gln	
		370					375					380					
	Asp	Ala	Tyr	Asn	Glu	Trp	Asn	Tyr	Ser	Lys	Phe	Arg	Leu	Trp	Gln	Leu	
	385					390					395					400	
45	Thr	Asp	Tyr	Asp	Lys	Ile	Ile	Phe	Ile	Asp	Ala	Asp	Leu	Leu	Ile	Leu	
					405					410					415		
	Arg	Asn	Ile	Asp	Phe	Leu	Phe	Glu	Met	Pro	Glu	Ile	Thr	Ala	Ile	Gly	
50				420					425					430			
	Asn	Asn	Ala	Thr	Leu	Phe	Asn	Ser	Gly	Val	Met	Val	Val	Glu	Pro	Ser	
			435					440					445				
55	Asn	Cys	Thr	Phe	Gln	Leu	Leu	Met	Asp	His	Ile	Asn	Glu	Ile	Glu	Ser	
		450					455					460					
	Tyr	Asn	Gly	Gly	Asp	Gln	Gly	Tyr	Leu	Asn	Glu	Ile	Phe	Thr	Trp	Trp	
	465					470					475					480	
60	His	Arg	Ile	Pro	Lys	His	Met	Asn	Phe	Leu	Lys	His	Tyr	Trp	Glu	Gly	
					485					490					495		

Asp Glu Glu Glu Lys Lys Gln Met Lys Thr Arg Leu Phe Gly Ala Asp
 500 505 510
 5 Pro Pro Val Leu Tyr Val Leu His Tyr Leu Gly Leu Lys Pro Trp Leu
 515 520 525
 Cys Phe Arg Asp Tyr Asp Cys Asn Trp Asn Val Gly Lys Leu Gln Glu
 530 535 540
 10 Phe Ala Ser Asp Val Ala His Arg Thr Trp Trp Lys Val His Asp Ala
 545 550 555 560
 Met Pro Asp Asn Leu His Lys Tyr Cys Leu Leu Arg Ser Lys Gln Lys
 565 570 575
 Ala Ala Leu Glu Trp Asp Arg Arg Glu Ala Glu Lys Ala Asn Phe Ser
 580 585 590
 20 Asp Gly His Trp Lys Ile Lys Ile Lys Asp Pro Arg Leu Glu Thr Cys
 595 600 605
 Tyr Glu Glu Phe Cys Phe Trp Glu Ser Met Leu Trp His Trp Gly Glu
 610 615 620
 25 Thr Asn Trp Thr Asp Asn Ala Thr Ser Ser Pro Thr Pro Pro Met Val
 625 630 635 640
 Asn Thr Ala Ser Leu Ser Ser Leu
 645
 30
 35 <210> 5
 <211> 1485
 <212> DNA
 <213> Arabidopsis thaliana
 40 <220>
 <221> CDS
 <222> (1)..(1485)
 <400> 5
 45 atg gat tta caa aga act ttg atg ttc tct tgt tgg gtt ctg tct ctt 48
 Met Asp Leu Gln Arg Thr Leu Met Phe Ser Cys Trp Val Leu Ser Leu
 1 5 10 15
 ttg atc atc aaa acg aca gcg tat aac gag aaa cag ctg ttc cag ccg 96
 Leu Ile Ile Lys Thr Thr Ala Tyr Asn Glu Lys Gln Leu Phe Gln Pro
 20 25 30
 50 ctt gaa acg gaa aac gca aac gcg atg acc gcg gtt atg gag cga gga 144
 Leu Glu Thr Glu Asn Ala Asn Ala Met Thr Ala Val Met Glu Arg Gly
 35 40 45
 55 tta aag acg cag cgg cgg ccg gag cac aag aac gct tat gcg acg atg 192
 Leu Lys Thr Gln Arg Arg Pro Glu His Lys Asn Ala Tyr Ala Thr Met
 50 55 60
 60 atg tac atg gga aca cca aga gac tac gag ttc tac gtt gcg aca cgt 240
 Met Tyr Met Gly Thr Pro Arg Asp Tyr Glu Phe Tyr Val Ala Thr Arg

	65	70	75	80	
5	gtc ttg atc aga tcg ctt aag agt ctc cac gtg gac gct gat atc gtc Val Leu Ile Arg Ser Leu Lys Ser Leu His Val Asp Ala Asp Ile Val	85	90	95	288
10	gtt ata gcc tcc ctc gac gtt cct atc aac tgg att cac gct ctg gaa Val Ile Ala Ser Leu Asp Val Pro Ile Asn Trp Ile His Ala Leu Glu	100	105	110	336
15	gaa gaa gat gga gct aaa gta gtg aga gta gag aat ctt gag aat cca Glu Glu Asp Gly Ala Lys Val Val Arg Val Glu Asn Leu Glu Asn Pro	115	120	125	384
20	tac aag aaa caa acc aac ttc gac aac aga ttc aag ctt agt cta aac Tyr Lys Lys Gln Thr Asn Phe Asp Asn Arg Phe Lys Leu Ser Leu Asn	130	135	140	432
25	aag ctc tac gct tgg tct ctc tct gat tat gac cgt gtt gta atg ctt Lys Leu Tyr Ala Trp Ser Leu Ser Asp Tyr Asp Arg Val Val Met Leu	145	150	155	480
30	gat gtc gac aat ctc ttt ctc aag aac acc gac gag ctc ttc cag tgt Asp Val Asp Asn Leu Phe Leu Lys Asn Thr Asp Glu Leu Phe Gln Cys	165	170	175	528
35	ggc caa ttt tgt gct gtc ttc atc aac cct tgc atc ttc cac act ggt Gly Gln Phe Cys Ala Val Phe Ile Asn Pro Cys Ile Phe His Thr Gly	180	185	190	576
40	ctc ttt gtg ttg cag cca tca atg gag gtc ttt aga gac atg ctt cat Leu Phe Val Leu Gln Pro Ser Met Glu Val Phe Arg Asp Met Leu His	195	200	205	624
45	gag ctt gaa gta aag aga gat aac cct gat gga gct gat caa ggc ttt Glu Leu Glu Val Lys Arg Asp Asn Pro Asp Gly Ala Asp Gln Gly Phe	210	215	220	672
50	ctt gtc agc tac ttc tct gat tta ctc aat cag cct ctc ttt cgt cct Leu Val Ser Tyr Phe Ser Asp Leu Leu Asn Gln Pro Leu Phe Arg Pro	225	230	235	720
55	cct ccc gat aac cgc acc gcg ctt aag gga cat ttt agg ctt cct ttg Pro Pro Asp Asn Arg Thr Ala Leu Lys Gly His Phe Arg Leu Pro Leu	245	250	255	768
60	gga tat caa atg gac gca tct tat tac tac ctt aag ctc aga tgg aac Gly Tyr Gln Met Asp Ala Ser Tyr Tyr Tyr Leu Lys Leu Arg Trp Asn	260	265	270	816
65	gta cca tgt gga cca aac agt gtg ata acg ttc cca gga gca gta tgg Val Pro Cys Gly Pro Asn Ser Val Ile Thr Phe Pro Gly Ala Val Trp	275	280	285	864
70	tta aag cca tgg tat tgg tgg tca tgg cct gtt ctt cct tta ggc ctt Leu Lys Pro Trp Tyr Trp Trp Ser Trp Pro Val Leu Pro Leu Gly Leu	290	295	300	912
75	tca tgg cac cac caa cgc cgc tac acg att agt tat tca gca gag atg Ser Trp His His Gln Arg Arg Tyr Thr Ile Ser Tyr Ser Ala Glu Met	305	310	315	960

	cct tgg gtc cta acc caa gca gtg ttc tac cta gga att ata cta gtc	1008
	Pro Trp Val Leu Thr Gln Ala Val Phe Tyr Leu Gly Ile Ile Leu Val	
	325 330 335	
5	aca cgt cta gcg cgt ccc aac atg acc aag cta tgt tac cga cgt tct	1056
	Thr Arg Leu Ala Arg Pro Asn Met Thr Lys Leu Cys Tyr Arg Arg Ser	
	340 345 350	
10	gat aag aat cta agc atg atc cag aca gct ttc aag ttt gtt gca ctc	1104
	Asp Lys Asn Leu Ser Met Ile Gln Thr Ala Phe Lys Phe Val Ala Leu	
	355 360 365	
15	ctc ttt atc ctc tca gcc tac att ata cca ttc ttc atc atc cca cag	1152
	Leu Phe Ile Leu Ser Ala Tyr Ile Ile Pro Phe Phe Ile Ile Pro Gln	
	370 375 380	
20	acg atc cac ccg ctc att ggt tgg tct ctc tac tta acc ggc tcc ttt	1200
	Thr Ile His Pro Leu Ile Gly Trp Ser Leu Tyr Leu Thr Gly Ser Phe	
	385 390 395 400	
25	gct ctc tct acc ata ccc atc aac gcc ttc ttg ctt ccc att ctc cct	1248
	Ala Leu Ser Thr Ile Pro Ile Asn Ala Phe Leu Leu Pro Ile Leu Pro	
	405 410 415	
30	gtc ata aca ccg tgg ctt ggc att ttc ggg aca ctc ctc gtg atg gct	1296
	Val Ile Thr Pro Trp Leu Gly Ile Phe Gly Thr Leu Leu Val Met Ala	
	420 425 430	
35	ttt cct tct tat cct gat ggc gtt gtt aga gca ttg tcg gtt ttc ggg	1344
	Phe Pro Ser Tyr Pro Asp Gly Val Val Arg Ala Leu Ser Val Phe Gly	
	435 440 445	
40	tat gca ttt tgt tgt gca ccg ttt cta tgg gtc tcc ttt gtg aag atc	1392
	Tyr Ala Phe Cys Cys Ala Pro Phe Leu Trp Val Ser Phe Val Lys Ile	
	450 455 460	
45	aca tcg cat ctt cag att atg att gac aaa gag gtt ttg ttt ccg cgg	1440
	Thr Ser His Leu Gln Ile Met Ile Asp Lys Glu Val Leu Phe Pro Arg	
	465 470 475 480	
50	ttg ggt gaa tcc gga gtc act tct ggt ctc agc aaa ttg tac tga	1485
	Leu Gly Glu Ser Gly Val Thr Ser Gly Leu Ser Lys Leu Tyr	
	485 490 495	
	<210> 6	
	<211> 494	
	<212> PRT	
	<213> Arabidopsis thaliana	
	<400> 6	
	Met Asp Leu Gln Arg Thr Leu Met Phe Ser Cys Trp Val Leu Ser Leu	
	1 5 10 15	
55	Leu Ile Ile Lys Thr Thr Ala Tyr Asn Glu Lys Gln Leu Phe Gln Pro	
	20 25 30	
	Leu Glu Thr Glu Asn Ala Asn Ala Met Thr Ala Val Met Glu Arg Gly	
	35 40 45	
60	Leu Lys Thr Gln Arg Arg Pro Glu His Lys Asn Ala Tyr Ala Thr Met	
	50 55 60	
	Met Tyr Met Gly Thr Pro Arg Asp Tyr Glu Phe Tyr Val Ala Thr Arg	

60 <213> Arabidopsis thaliana

14

<220>

<221> CDS

<222> (1)..(1494)

5 <400> 7
 atg gat ttg cag aga ggc ttt gtg ttc ttg tct ttg gtt cta tct ttt 48
 Met Asp Leu Gln Arg Gly Phe Val Phe Leu Ser Leu Val Leu Ser Phe
 1 5 10 15

10 atg ata atc gaa acg acg gcg tat cga gag aga cag ctg ctg ctg ctg 96
 Met Ile Ile Glu Thr Thr Ala Tyr Arg Glu Arg Gln Leu Leu Leu Leu
 20 25 30

15 caa cca ccg caa gaa acg gcg ata gat acc gca aac gcg gtg gtg acg 144
 Gln Pro Pro Gln Glu Thr Ala Ile Asp Thr Ala Asn Ala Val Val Thr
 35 40 45

20 gtt caa gat cga ggt ttg aag acg cgg cga ccg gag cat aag aac gca 192
 Val Gln Asp Arg Gly Leu Lys Thr Arg Arg Pro Glu His Lys Asn Ala
 50 55 60

25 tac gca acg atg atg tac atg ggg acg cca aga gac tac gag ttc tac 240
 Tyr Ala Thr Met Met Tyr Met Gly Thr Pro Arg Asp Tyr Glu Phe Tyr
 65 70 75 80

30 gtt gcg aca cgt gtt ttg atc aga tcg ttg aga agt ctt cac gtg gaa 288
 Val Ala Thr Arg Val Leu Ile Arg Ser Leu Arg Ser Leu His Val Glu
 85 90 95

35 gct gat ctc gtc gtc atc gct tct ctc gac gtt cct ctc cga tgg gtt 336
 Ala Asp Leu Val Val Ile Ala Ser Leu Asp Val Pro Leu Arg Trp Val
 100 105 110

40 caa acc ttg gaa gag gaa gat gga gct aaa gtg gtg aga gtt gaa aat 384
 Gln Thr Leu Glu Glu Glu Asp Gly Ala Lys Val Val Arg Val Glu Asn
 115 120 125

45 gtg gat aat cca tac agg aga cag acc aac ttc aac agt aga ttc aag 432
 Val Asp Asn Pro Tyr Arg Arg Gln Thr Asn Phe Asn Ser Arg Phe Lys
 130 135 140

50 ctt act cta aac aag ctc tac gct tgg gct ttg tct gat tac gac cgt 480
 Leu Thr Leu Asn Lys Leu Tyr Ala Trp Ala Leu Ser Asp Tyr Asp Arg
 145 150 155 160

55 gtg gtc atg cta gat gcc gat aac ctc ttt ctt aag aaa gcc gac gag 528
 Val Val Met Leu Asp Ala Asp Asn Leu Phe Leu Lys Lys Ala Asp Glu
 165 170 175

60 ttg ttc cag tgt ggg cgc ttc tgt gcg gtc ttc att aac cct tgt atc 576
 Leu Phe Gln Cys Gly Arg Phe Cys Ala Val Phe Ile Asn Pro Cys Ile
 180 185 190

65 ttc cac act ggt ctc ttc gtg ttg cag cca tca gtg gaa gtg ttc aag 624
 Phe His Thr Gly Leu Phe Val Leu Gln Pro Ser Val Glu Val Phe Lys
 195 200 205

70 gac atg ctc cat gag cta caa gtt gga aga aag aat cct gat gga gct 672
 Asp Met Leu His Glu Leu Gln Val Gly Arg Lys Asn Pro Asp Gly Ala
 210 215 220

15

	gat	caa	ggt	ttt	ctt	gtc	agt	tac	ttc	tct	gat	ctt	ctt	gac	caa	cct	720
	Asp	Gln	Gly	Phe	Leu	Val	Ser	Tyr	Phe	Ser	Asp	Leu	Leu	Asp	Gln	Pro	
	225					230					235					240	
5	ctc	ttt	agt	cct	ccg	agt	aac	gga	tct	gta	ctt	aat	ggt	cac	ttg	aga	768
	Leu	Phe	Ser	Pro	Pro	Ser	Asn	Gly	Ser	Val	Leu	Asn	Gly	His	Leu	Arg	
					245					250						255	
10	ctt	ccc	tta	ggc	tac	caa	atg	gac	gct	tct	tat	ttc	tat	ctt	aag	cta	816
	Leu	Pro	Leu	Gly	Tyr	Gln	Met	Asp	Ala	Ser	Tyr	Phe	Tyr	Leu	Lys	Leu	
				260					265							270	
15	aga	tgg	aac	ata	ccc	tgt	gga	cca	aac	agt	gtg	att	aca	ttc	ccg	gga	864
	Arg	Trp	Asn	Ile	Pro	Cys	Gly	Pro	Asn	Ser	Val	Ile	Thr	Phe	Pro	Gly	
			275					280									
20	gct	gtt	tgg	tta	aag	cca	tgg	tac	tgg	tgg	tca	tgg	cct	gtt	ctt	cca	912
	Ala	Val	Trp	Leu	Lys	Pro	Trp	Tyr	Trp	Trp	Ser	Trp	Pro	Val	Leu	Pro	
		290					295					300					
25	cta	ggt	ttc	tca	tgg	cac	gag	cag	cgt	cgc	gcc	act	ata	ggg	tac	tca	960
	Leu	Gly	Phe	Ser	Trp	His	Glu	Gln	Arg	Arg	Ala	Thr	Ile	Gly	Tyr	Ser	
	305					310					315					320	
30	gcc	gaa	atg	cct	ttg	gtt	ata	atc	caa	gca	atg	ttt	tac	ctt	gga	atc	1008
	Ala	Glu	Met	Pro	Leu	Val	Ile	Ile	Gln	Ala	Met	Phe	Tyr	Leu	Gly	Ile	
					325					330					335		
35	ata	gtg	gtt	aca	cga	cta	gct	cgt	cct	aac	ata	acc	aag	cta	tgt	tat	1056
	Ile	Val	Val	Thr	Arg	Leu	Ala	Arg	Pro	Asn	Ile	Thr	Lys	Leu	Cys	Tyr	
				340					345					350			
40	cgc	cgc	tct	gac	cgc	aac	tta	aca	acg	atc	caa	gct	ggt	ttt	aag	ttg	1104
	Arg	Arg	Ser	Asp	Arg	Asn	Leu	Thr	Thr	Ile	Gln	Ala	Gly	Phe	Lys	Leu	
			355					360					365				
45	atc	gcg	ctt	ctc	tct	gta	gtt	gca	gcc	tac	atc	ttc	cca	ttc	ttc	acc	1152
	Ile	Ala	Leu	Leu	Ser	Val	Val	Ala	Ala	Tyr	Ile	Phe	Pro	Phe	Phe	Thr	
		370					375					380					
50	atc	cct	cac	act	atc	cac	cca	ctc	atc	ggc	tgg	tcg	ctc	tac	ttg	atg	1200
	Ile	Pro	His	Thr	Ile	His	Pro	Leu	Ile	Gly	Trp	Ser	Leu	Tyr	Leu	Met	
		385				390				395						400	
55	gct	tct	ttt	gct	ctc	tct	tcc	att	tca	atc	aac	act	ctc	ctc	ctc	cca	1248
	Ala	Ser	Phe	Ala	Leu	Ser	Ser	Ile	Ser	Ile	Asn	Thr	Leu	Leu	Leu	Pro	
					405					410						415	
60	acg	ctc	cct	gtt	ctc	act	cca	tgg	cta	gga	att	ctc	ggc	act	ctc	ctt	1296
	Thr	Leu	Pro	Val	Leu	Thr	Pro	Trp	Leu	Gly	Ile	Leu	Gly	Thr	Leu	Leu	
				420					425					430			
65	gtc	atg	gcc	ttc	cct	tgg	tac	cct	gat	gga	gtg	gtc	aga	gcc	ttg	tca	1344
	Val	Met	Ala	Phe	Pro	Trp	Tyr	Pro	Asp	Gly	Val	Val	Arg	Ala	Leu	Ser	
			435					440					445				
70	gtt	ttc	gca	tac	gca	ttt	tgt	tgc	gca	ccc	ttt	gtg	tgg	gtt	tca	ttc	1392
	Val	Phe	Ala	Tyr	Ala	Phe	Cys	Cys	Ala	Pro	Phe	Val	Trp	Val	Ser	Phe	
		450					455					460					
75	cgc	aaa	atc	aca	tcg	cac	ctc	cag	gtt	ttg	att	gag	aaa	gag	gtg	ttg	1440

16

	Arg	Lys	Ile	Thr	Ser	His	Leu	Gln	Val	Leu	Ile	Glu	Lys	Glu	Val	Leu	
	465					470					475					480	
5	ttc	ccg	cga	ttg	gga	gac	tca	ggg	gtc	act	tca	ggc	ttc	agc	aaa	ttg	1488
	Phe	Pro	Arg	Leu	Gly	Asp	Ser	Gly	Val	Thr	Ser	Gly	Phe	Ser	Lys	Leu	
					485					490					495		
10	tat	tag															1494
	Tyr																
	<210> 8																
	<211> 497																
	<212> PRT																
15	<213> Arabidopsis thaliana																
	<400> 8																
	Met	Asp	Leu	Gln	Arg	Gly	Phe	Val	Phe	Leu	Ser	Leu	Val	Leu	Ser	Phe	
	1				5					10					15		
20	Met	Ile	Ile	Glu	Thr	Thr	Ala	Tyr	Arg	Glu	Arg	Gln	Leu	Leu	Leu	Leu	
				20					25					30			
	Gln	Pro	Pro	Gln	Glu	Thr	Ala	Ile	Asp	Thr	Ala	Asn	Ala	Val	Val	Thr	
			35					40					45				
25	Val	Gln	Asp	Arg	Gly	Leu	Lys	Thr	Arg	Arg	Pro	Glu	His	Lys	Asn	Ala	
	50						55					60					
	Tyr	Ala	Thr	Met	Met	Tyr	Met	Gly	Thr	Pro	Arg	Asp	Tyr	Glu	Phe	Tyr	
	65					70					75					80	
	Val	Ala	Thr	Arg	Val	Leu	Ile	Arg	Ser	Leu	Arg	Ser	Leu	His	Val	Glu	
					85					90					95		
30	Ala	Asp	Leu	Val	Val	Ile	Ala	Ser	Leu	Asp	Val	Pro	Leu	Arg	Trp	Val	
				100					105					110			
	Gln	Thr	Leu	Glu	Glu	Glu	Asp	Gly	Ala	Lys	Val	Val	Arg	Val	Glu	Asn	
			115					120					125				
35	Val	Asp	Asn	Pro	Tyr	Arg	Arg	Gln	Thr	Asn	Phe	Asn	Ser	Arg	Phe	Lys	
	130					135						140					
	Leu	Thr	Leu	Asn	Lys	Leu	Tyr	Ala	Trp	Ala	Leu	Ser	Asp	Tyr	Asp	Arg	
	145				150					155					160		
	Val	Val	Met	Leu	Asp	Ala	Asp	Asn	Leu	Phe	Leu	Lys	Lys	Ala	Asp	Glu	
					165					170					175		
40	Leu	Phe	Gln	Cys	Gly	Arg	Phe	Cys	Ala	Val	Phe	Ile	Asn	Pro	Cys	Ile	
				180					185					190			
	Phe	His	Thr	Gly	Leu	Phe	Val	Leu	Gln	Pro	Ser	Val	Glu	Val	Phe	Lys	
			195					200					205				
45	Asp	Met	Leu	His	Glu	Leu	Gln	Val	Gly	Arg	Lys	Asn	Pro	Asp	Gly	Ala	
	210					215						220					
	Asp	Gln	Gly	Phe	Leu	Val	Ser	Tyr	Phe	Ser	Asp	Leu	Leu	Asp	Gln	Pro	
	225					230					235				240		
	Leu	Phe	Ser	Pro	Pro	Ser	Asn	Gly	Ser	Val	Leu	Asn	Gly	His	Leu	Arg	
					245					250					255		
50	Leu	Pro	Leu	Gly	Tyr	Gln	Met	Asp	Ala	Ser	Tyr	Phe	Tyr	Leu	Lys	Leu	
				260					265					270			
	Arg	Trp	Asn	Ile	Pro	Cys	Gly	Pro	Asn	Ser	Val	Ile	Thr	Phe	Pro	Gly	
			275					280					285				
55	Ala	Val	Trp	Leu	Lys	Pro	Trp	Tyr	Trp	Trp	Ser	Trp	Pro	Val	Leu	Pro	
	290					295						300					
	Leu	Gly	Phe	Ser	Trp	His	Glu	Gln	Arg	Arg	Ala	Thr	Ile	Gly	Tyr	Ser	
	305					310					315				320		
	Ala	Glu	Met	Pro	Leu	Val	Ile	Ile	Gln	Ala	Met	Phe	Tyr	Leu	Gly	Ile	
					325					330					335		
60	Ile	Val	Val	Thr	Arg	Leu	Ala	Arg	Pro	Asn	Ile	Thr	Lys	Leu	Cys	Tyr	
				340					345					350			

17

	Arg	Arg	Ser	Asp	Arg	Asn	Leu	Thr	Thr	Ile	Gln	Ala	Gly	Phe	Lys	Leu	
			355					360					365				
	Ile	Ala	Leu	Leu	Ser	Val	Val	Ala	Ala	Tyr	Ile	Phe	Pro	Phe	Phe	Thr	
		370					375					380					
5	Ile	Pro	His	Thr	Ile	His	Pro	Leu	Ile	Gly	Trp	Ser	Leu	Tyr	Leu	Met	
	385					390					395					400	
	Ala	Ser	Phe	Ala	Leu	Ser	Ser	Ile	Ser	Ile	Asn	Thr	Leu	Leu	Leu	Pro	
				405						410						415	
10	Thr	Leu	Pro	Val	Leu	Thr	Pro	Trp	Leu	Gly	Ile	Leu	Gly	Thr	Leu	Leu	
				420					425					430			
	Val	Met	Ala	Phe	Pro	Trp	Tyr	Pro	Asp	Gly	Val	Val	Arg	Ala	Leu	Ser	
			435					440					445				
	Val	Phe	Ala	Tyr	Ala	Phe	Cys	Cys	Ala	Pro	Phe	Val	Trp	Val	Ser	Phe	
		450					455					460					
15	Arg	Lys	Ile	Thr	Ser	His	Leu	Gln	Val	Leu	Ile	Glu	Lys	Glu	Val	Leu	
	465					470					475					480	
	Phe	Pro	Arg	Leu	Gly	Asp	Ser	Gly	Val	Thr	Ser	Gly	Phe	Ser	Lys	Leu	
				485						490					495		
20	Tyr																
	<210>	9															
	<211>	1944															
25	<212>	DNA															
	<213>	Arabidopsis thaliana															
	<220>																
	<221>	CDS															
30	<222>	(1)..(1944)															
	<400>	9															
	atg	cat	cca	gct	tca	tgt	agt	ctc	tca	ctc	agg	aaa	gtc	aaa	ctt	aat	48
35	Met	His	Pro	Ala	Ser	Cys	Ser	Leu	Ser	Leu	Arg	Lys	Val	Lys	Leu	Asn	
	1				5					10					15		
	att	ctt	aca	gtg	aga	atg	aag	ctc	tct	tct	gaa	agt	ccc	atg	gcg	ccg	96
	Ile	Leu	Thr	Val	Arg	Met	Lys	Leu	Ser	Ser	Glu	Ser	Pro	Met	Ala	Pro	
				20					25					30			
40	tca	agt	cag	tca	agt	cat	cga	ctt	tac	att	tcc	agt	gag	aaa	aca	aaa	144
	Ser	Ser	Gln	Ser	Ser	His	Arg	Leu	Tyr	Ile	Ser	Ser	Glu	Lys	Thr	Lys	
			35					40					45				
45	acg	aag	aga	ttc	caa	aga	aac	gga	tac	act	ctc	gat	gtt	gaa	atg	tgt	192
	Thr	Lys	Arg	Phe	Gln	Arg	Asn	Gly	Tyr	Thr	Leu	Asp	Val	Glu	Met	Cys	
		50					55					60					
	gtc	aac	ttc	tct	agt	ctg	aaa	ctt	gtt	ttg	ttt	ctt	atg	atg	ctg	gtt	240
50	Val	Asn	Phe	Ser	Ser	Leu	Lys	Leu	Val	Leu	Phe	Leu	Met	Met	Leu	Val	
	65					70				75					80		
	gct	atg	ttc	aca	ctc	tac	tgt	tct	cca	ccg	ttg	caa	att	cct	gaa	gat	288
55	Ala	Met	Phe	Thr	Leu	Tyr	Cys	Ser	Pro	Pro	Leu	Gln	Ile	Pro	Glu	Asp	
				85					90					95			
	cca	tca	agt	ttt	gca	aac	aaa	tgg	ata	cta	gaa	cct	gct	gta	acc	aca	336
	Pro	Ser	Ser	Phe	Ala	Asn	Lys	Trp	Ile	Leu	Glu	Pro	Ala	Val	Thr	Thr	
				100				105					110				
60	gat	cct	cgc	tat	ata	gct	aca	tct	gag	atc	aac	tgg	aac	agt	atg	tca	384

18

	Asp	Pro	Arg	Tyr	Ile	Ala	Thr	Ser	Glu	Ile	Asn	Trp	Asn	Ser	Met	Ser	
			115					120					125				
5	ctt	gtt	gtt	gag	cat	tac	tta	tct	ggc	aga	agc	gag	tat	caa	gga	att	432
	Leu	Val	Val	Glu	His	Tyr	Leu	Ser	Gly	Arg	Ser	Glu	Tyr	Gln	Gly	Ile	
		130					135					140					
10	ggc	ttt	cta	aat	ctc	aac	gat	aac	gag	att	aat	cga	tgg	cag	gtg	gtc	480
	Gly	Phe	Leu	Asn	Leu	Asn	Asp	Asn	Glu	Ile	Asn	Arg	Trp	Gln	Val	Val	
	145					150					155					160	
15	ata	aaa	tct	cac	tgt	cag	cat	ata	gct	ttg	cat	cta	gac	cat	gct	gca	528
	Ile	Lys	Ser	His	Cys	Gln	His	Ile	Ala	Leu	His	Leu	Asp	His	Ala	Ala	
					165					170					175		
20	agt	aac	ata	act	tgg	aaa	tct	tta	tac	ccg	gaa	tgg	att	gac	gag	gaa	576
	Ser	Asn	Ile	Thr	Trp	Lys	Ser	Leu	Tyr	Pro	Glu	Trp	Ile	Asp	Glu	Glu	
				180					185					190			
25	gaa	aaa	ttc	aaa	gtc	ccc	act	tgt	cct	tct	ctt	cct	tgg	att	caa	gtt	624
	Glu	Lys	Phe	Lys	Val	Pro	Thr	Cys	Pro	Ser	Leu	Pro	Trp	Ile	Gln	Val	
			195					200					205				
30	cct	gac	aag	tct	cga	atc	gat	ctt	atc	att	gcc	aag	ctc	cca	tgt	aac	672
	Pro	Asp	Lys	Ser	Arg	Ile	Asp	Leu	Ile	Ile	Ala	Lys	Leu	Pro	Cys	Asn	
		210					215					220					
35	aag	tca	gga	aaa	tgg	tca	aga	gat	gtg	gct	aga	ttg	cac	tta	caa	ctt	720
	Lys	Ser	Gly	Lys	Trp	Ser	Arg	Asp	Val	Ala	Arg	Leu	His	Leu	Gln	Leu	
	225					230					235					240	
40	gca	gca	gct	cga	gtg	gcg	gca	tct	tct	gaa	ggg	ctt	cat	gat	gtt	cat	768
	Ala	Ala	Ala	Arg	Val	Ala	Ala	Ser	Ser	Glu	Gly	Leu	His	Asp	Val	His	
					245					250					255		
45	gtg	att	ttg	gta	tca	gat	tgc	ttt	cca	ata	ccg	aat	ctt	ttt	acg	ggc	816
	Val	Ile	Leu	Val	Ser	Asp	Cys	Phe	Pro	Ile	Pro	Asn	Leu	Phe	Thr	Gly	
				260					265					270			
50	caa	gaa	ctt	gtt	gcc	cgt	caa	gga	aac	ata	tgg	ctg	tat	aag	cct	aaa	864
	Gln	Glu	Leu	Val	Ala	Arg	Gln	Gly	Asn	Ile	Trp	Leu	Tyr	Lys	Pro	Lys	
			275					280					285				
55	ctt	cac	cag	tta	aga	caa	aag	tta	caa	ctt	cct	gtt	ggc	tcc	tgt	gaa	912
	Leu	His	Gln	Leu	Arg	Gln	Lys	Leu	Gln	Leu	Pro	Val	Gly	Ser	Cys	Glu	
		290					295					300					
60	ctt	tct	gtt	cct	ctt	caa	gct	aaa	gat	aat	ttc	tac	tcg	gca	aat	gcc	960
	Leu	Ser	Val	Pro	Leu	Gln	Ala	Lys	Asp	Asn	Phe	Tyr	Ser	Ala	Asn	Ala	
	305					310					315					320	
65	aag	aaa	gaa	gcg	tac	gcg	acg	atc	ttg	cac	tca	gat	gat	gct	ttt	gtc	1008
	Lys	Lys	Glu	Ala	Tyr	Ala	Thr	Ile	Leu	His	Ser	Asp	Asp	Ala	Phe	Val	
					325					330					335		
70	tgt	gga	gcc	att	gca	gta	gca	cag	agc	att	cga	atg	tca	ggc	tct	act	1056
	Cys	Gly	Ala	Ile	Ala	Val	Ala	Gln	Ser	Ile	Arg	Met	Ser	Gly	Ser	Thr	
				340					345					350			
75	cgc	aat	ttg	gta	ata	cta	gtc	gat	gat	tcg	atc	agt	gaa	tac	cat	aga	1104
	Arg	Asn	Leu	Val	Ile	Leu	Val	Asp	Asp	Ser	Ile	Ser	Glu	Tyr	His	Arg	

	355	360	365	
5	agt ggc ttg gaa tca gct gga tgg aag att cac aca ttt caa aga atc Ser Gly Leu Glu Ser Ala Gly Trp Lys Ile His Thr Phe Gln Arg Ile 370 375 380	1152		
10	aga aac ccg aaa gct gaa gca aat gca tat aac caa tgg aac tac agc Arg Asn Pro Lys Ala Glu Ala Asn Ala Tyr Asn Gln Trp Asn Tyr Ser 385 390 395 400	1200		
15	aaa ttc cgt ctt tgg gaa ttg aca gaa tac aac aag atc atc ttc att Lys Phe Arg Leu Trp Glu Leu Thr Glu Tyr Asn Lys Ile Ile Phe Ile 405 410 415	1248		
20	gat gca gac atg ctt atc ctc aga aac atg gat ttc ctc ttc gag tac Asp Ala Asp Met Leu Ile Leu Arg Asn Met Asp Phe Leu Phe Glu Tyr 420 425 430	1296		
25	ccc gaa atc tcc aca act gga aac gac ggt acg ctc ttc aac tcc ggt Pro Glu Ile Ser Thr Thr Gly Asn Asp Gly Thr Leu Phe Asn Ser Gly 435 440 445	1344		
30	cta atg gtg att gaa cca tca aat tca aca ttc cag tta cta atg gat Leu Met Val Ile Glu Pro Ser Asn Ser Thr Phe Gln Leu Leu Met Asp 450 455 460	1392		
35	cac atc aac gat atc aat tcc tac aat gga gga gac caa ggt tac ctt His Ile Asn Asp Ile Asn Ser Tyr Asn Gly Gly Asp Gln Gly Tyr Leu 465 470 475 480	1440		
40	aac gag ata ttc aca tgg tgg cat cgg att cca aaa cac atg aat ttc Asn Glu Ile Phe Thr Trp Trp His Arg Ile Pro Lys His Met Asn Phe 485 490 495	1488		
45	ttg aag cat ttc tgg gaa gga gac aca cct aag cac agg aaa tct aag Leu Lys His Phe Trp Glu Gly Asp Thr Pro Lys His Arg Lys Ser Lys 500 505 510	1536		
50	acg aga cta ttt gga gct gat cct ccg ata ctc tac gtt ctt cat tac Thr Arg Leu Phe Gly Ala Asp Pro Pro Ile Leu Tyr Val Leu His Tyr 515 520 525	1584		
55	cta ggt tac aac aaa cca tgg gta tgc ttc aga gac tac gat tgc aat Leu Gly Tyr Asn Lys Pro Trp Val Cys Phe Arg Asp Tyr Asp Cys Asn 530 535 540	1632		
60	tgg aat gtc gtt gga tac cat caa ttc gcg agc gat gaa gca cac aaa Trp Asn Val Val Gly Tyr His Gln Phe Ala Ser Asp Glu Ala His Lys 545 550 555 560	1680		
65	act tgg tgg aga gtg cac gac gcg atg cct aag aaa ttg cag agg ttt Thr Trp Trp Arg Val His Asp Ala Met Pro Lys Lys Leu Gln Arg Phe 565 570 575	1728		
70	tgt cta ctg agt tcg aaa caa aag gcg caa ctt gag tgg gat cgg aga Cys Leu Leu Ser Ser Lys Gln Lys Ala Gln Leu Glu Trp Asp Arg Arg 580 585 590	1776		
75	caa gct gag aaa gcg aat tac aga gac gga cat tgg agg att aag atc Gln Ala Glu Lys Ala Asn Tyr Arg Asp Gly His Trp Arg Ile Lys Ile 595 600 605	1824		

	aaa	gat	aag	aga	ctt	acg	act	tgt	ttt	gaa	gat	ttc	tgt	ttc	tgg	gag	1872
	Lys	Asp	Lys	Arg	Leu	Thr	Thr	Cys	Phe	Glu	Asp	Phe	Cys	Phe	Trp	Glu	
		610					615					620					
5	agt	atg	ctt	tgg	cat	tgg	ggg	gag	act	cag	acc	aac	tcc	acc	gtc	gct	1920
	Ser	Met	Leu	Trp	His	Trp	Gly	Glu	Thr	Gln	Thr	Asn	Ser	Thr	Val	Ala	
		625				630					635					640	
10	gct	gat	tct	tcc	tcc	acc	gcg	taa									1944
	Ala	Asp	Ser	Ser	Ser	Thr	Ala										
					645												
15	<210>	10															
	<211>	647															
	<212>	PRT															
	<213>	Arabidopsis thaliana															
20	<400>	10															
	Met	His	Pro	Ala	Ser	Cys	Ser	Leu	Ser	Leu	Arg	Lys	Val	Lys	Leu	Asn	
	1				5					10					15		
	Ile	Leu	Thr	Val	Arg	Met	Lys	Leu	Ser	Ser	Glu	Ser	Pro	Met	Ala	Pro	
				20					25					30			
25	Ser	Ser	Gln	Ser	Ser	His	Arg	Leu	Tyr	Ile	Ser	Ser	Glu	Lys	Thr	Lys	
			35					40					45				
	Thr	Lys	Arg	Phe	Gln	Arg	Asn	Gly	Tyr	Thr	Leu	Asp	Val	Glu	Met	Cys	
		50					55					60					
	Val	Asn	Phe	Ser	Ser	Leu	Lys	Leu	Val	Leu	Phe	Leu	Met	Met	Leu	Val	
30	65					70					75					80	
	Ala	Met	Phe	Thr	Leu	Tyr	Cys	Ser	Pro	Pro	Leu	Gln	Ile	Pro	Glu	Asp	
					85					90					95		
	Pro	Ser	Ser	Phe	Ala	Asn	Lys	Trp	Ile	Leu	Glu	Pro	Ala	Val	Thr	Thr	
				100					105					110			
35	Asp	Pro	Arg	Tyr	Ile	Ala	Thr	Ser	Glu	Ile	Asn	Trp	Asn	Ser	Met	Ser	
			115					120					125				
	Leu	Val	Val	Glu	His	Tyr	Leu	Ser	Gly	Arg	Ser	Glu	Tyr	Gln	Gly	Ile	
		130					135					140					
	Gly	Phe	Leu	Asn	Leu	Asn	Asp	Asn	Glu	Ile	Asn	Arg	Trp	Gln	Val	Val	
40	145					150					155					160	
	Ile	Lys	Ser	His	Cys	Gln	His	Ile	Ala	Leu	His	Leu	Asp	His	Ala	Ala	
				165					170						175		
	Ser	Asn	Ile	Thr	Trp	Lys	Ser	Leu	Tyr	Pro	Glu	Trp	Ile	Asp	Glu	Glu	
				180					185					190			
45	Glu	Lys	Phe	Lys	Val	Pro	Thr	Cys	Pro	Ser	Leu	Pro	Trp	Ile	Gln	Val	
			195					200					205				
	Pro	Asp	Lys	Ser	Arg	Ile											

21

						325					330					335					
					Cys	Gly	Ala	Ile	Ala	Val	Ala	Gln	Ser	Ile	Arg	Met	Ser	Gly	Ser	Thr	
								340					345					350			
5					Arg	Asn	Leu	Val	Ile	Leu	Val	Asp	Asp	Ser	Ile	Ser	Glu	Tyr	His	Arg	
								355				360					365				
					Ser	Gly	Leu	Glu	Ser	Ala	Gly	Trp	Lys	Ile	His	Thr	Phe	Gln	Arg	Ile	
								370				375				380					
					Arg	Asn	Pro	Lys	Ala	Glu	Ala	Asn	Ala	Tyr	Asn	Gln	Trp	Asn	Tyr	Ser	
										390					395					400	
10					Lys	Phe	Arg	Leu	Trp	Glu	Leu	Thr	Glu	Tyr	Asn	Lys	Ile	Ile	Phe	Ile	
								405						410					415		
					Asp	Ala	Asp	Met	Leu	Ile	Leu	Arg	Asn	Met	Asp	Phe	Leu	Phe	Glu	Tyr	
								420					425					430			
15					Pro	Glu	Ile	Ser	Thr	Thr	Gly	Asn	Asp	Gly	Thr	Leu	Phe	Asn	Ser	Gly	
								435				440					445				
					Leu	Met	Val	Ile	Glu	Pro	Ser	Asn	Ser	Thr	Phe	Gln	Leu	Leu	Met	Asp	
										455						460					
					His	Ile	Asn	Asp	Ile	Asn	Ser	Tyr	Asn	Gly	Gly	Asp	Gln	Gly	Tyr	Leu	
										470					475					480	
20					Asn	Glu	Ile	Phe	Thr	Trp	Trp	His	Arg	Ile	Pro	Lys	His	Met	Asn	Phe	
										485				490					495		
					Leu	Lys	His	Phe	Trp	Glu	Gly	Asp	Thr	Pro	Lys	His	Arg	Lys	Ser	Lys	
								500				505					510				
25					Thr	Arg	Leu	Phe	Gly	Ala	Asp	Pro	Pro	Ile	Leu	Tyr	Val	Leu	His	Tyr	
								515				520					525				
					Leu	Gly	Tyr	Asn	Lys	Pro	Trp	Val	Cys	Phe	Arg	Asp	Tyr	Asp	Cys	Asn	
											535					540					
					Trp	Asn	Val	Val	Gly	Tyr	His	Gln	Phe	Ala	Ser	Asp	Glu	Ala	His	Lys	
										550					555					560	
30					Thr	Trp	Trp	Arg	Val	His	Asp	Ala	Met	Pro	Lys	Lys	Leu	Gln	Arg	Phe	
									565					570					575		
					Cys	Leu	Leu	Ser	Ser	Lys	Gln	Lys	Ala	Gln	Leu	Glu	Trp	Asp	Arg	Arg	
								580				585					590				
35					Gln	Ala	Glu	Lys	Ala	Asn	Tyr	Arg	Asp	Gly	His	Trp	Arg	Ile	Lys	Ile	
								595				600				605					
					Lys	Asp	Lys	Arg	Leu	Thr	Thr	Cys	Phe	Glu	Asp	Phe	Cys	Phe	Trp	Glu	
										615					620						
					Ser	Met	Leu	Trp	His	Trp	Gly	Glu	Thr	Gln	Thr	Asn	Ser	Thr	Val	Ala	
										630					635					640	
40					Ala	Asp	Ser	Ser	Ser	Thr	Ala										
									645												
45					<210> 11																
					<211> 1857																
					<212> DNA																
					<213> Arabidopsis thaliana																
50					<220>																
					<221> CDS																
					<222> (1)..(1857)																
					<400> 11																
55					atg	ata	oct	tcc	tca	agt	ccc	atg	gag	tca	aga	cat	cga	ctc	tcg	ttc	48
					Met	Ile	Pro	Ser	Ser	Ser	Pro	Met	Glu	Ser	Arg	His	Arg	Leu	Ser	Phe	
					1				5				10					15			
					tca	aaa	gag	aag	aca	agt	agg	agg	aga	ttt	caa	aga	att	gag	aag	ggg	96
60					Ser	Lys	Glu	Lys	Thr	Ser	Arg	Arg	Arg	Phe	Gln	Arg	Ile	Glu	Lys	Gly	
							20					25					30				

5	gtc aag ttc aac act ctg aaa ctt gtg ttg att tgt ata atg ctt gga	144
	Val Lys Phe Asn Thr Leu Lys Leu Val Leu Ile Cys Ile Met Leu Gly	
	35 40 45	
	gct ttg ttc acg atc tac cgt ttt cgt tat cca ccg cta caa att cct	192
	Ala Leu Phe Thr Ile Tyr Arg Phe Arg Tyr Pro Pro Leu Gln Ile Pro	
10	50 55 60	
	gaa att cca act agt ttt ggt ctt act act gat cct cgc tat gta gct	240
	Glu Ile Pro Thr Ser Phe Gly Leu Thr Thr Asp Pro Arg Tyr Val Ala	
	65 70 75 80	
	aca gct gag atc aac tgg aac cat atg tca aat ctt gtt gag aag cac	288
15	Thr Ala Glu Ile Asn Trp Asn His Met Ser Asn Leu Val Glu Lys His	
	85 90 95	
	gta ttt ggt aga agc gag tat caa gga att ggt ctt ata aat ctt aac	336
	Val Phe Gly Arg Ser Glu Tyr Gln Gly Ile Gly Leu Ile Asn Leu Asn	
	100 105 110	
20	gat aac gag att gat cga ttc aag gag gta acg aaa tct gac tgt gat	384
	Asp Asn Glu Ile Asp Arg Phe Lys Glu Val Thr Lys Ser Asp Cys Asp	
	115 120 125	
	cat gta gct ttg cat cta gat tat gct gca aag aac ata aca tgg gaa	432
	His Val Ala Leu His Leu Asp Tyr Ala Ala Lys Asn Ile Thr Trp Glu	
25	130 135 140	
	tct tta tac ccg gaa tgg att gat gaa gtt gaa gaa ttc gaa gtc cct	480
	Ser Leu Tyr Pro Glu Trp Ile Asp Glu Val Glu Glu Phe Glu Val Pro	
	145 150 155 160	
	act tgt cct tct ctg cct ttg att caa att cct ggc aag cct cgg att	528
30	Thr Cys Pro Ser Leu Pro Leu Ile Gln Ile Pro Gly Lys Pro Arg Ile	
	165 170 175	
	gat ctt gta att gcc aag ctt ccg tgt gat aaa tca gga aaa tgg tct	576
	Asp Leu Val Ile Ala Lys Leu Pro Cys Asp Lys Ser Gly Lys Trp Ser	
	180 185 190	
35	aga gat gtg gct cgc ttg cat tta caa ctt gca gca gct cga gtg gcg	624
	Arg Asp Val Ala Arg Leu His Leu Gln Leu Ala Ala Arg Val Ala	
	195 200 205	
	gct tct tct aaa gga ctt cat aat gtt cat gtg att ttg gta tct gat	672
	Ala Ser Ser Lys Gly Leu His Asn Val His Val Ile Leu Val Ser Asp	
40	210 215 220	
	tgc ttt cca ata ccg aat ctt ttt acg ggt caa gaa ctt gtt gcc cgt	720
	Cys Phe Pro Ile Pro Asn Leu Phe Thr Gly Gln Glu Leu Val Ala Arg	
	225 230 235 240	
	caa gga aac ata tgg ctg tat aag cct aat ctt cac cag cta aga caa	768
45	Gln Gly Asn Ile Trp Leu Tyr Lys Pro Asn Leu His Gln Leu Arg Gln	
	245 250 255	
	aag tta cag ctt cct gtt ggt tcc tgt gaa ctt tct gtt cct ctt caa	816
	Lys Leu Gln Leu Pro Val Gly Ser Cys Glu Leu Ser Val Pro Leu Gln	
	260 265 270	
50	60	

23

	gct	aaa	gat	aat	ttc	tac	tcc	gca	ggt	gca	aag	aaa	gaa	gct	tac	gcg	864
	Ala	Lys	Asp	Asn	Phe	Tyr	Ser	Ala	Gly	Ala	Lys	Lys	Glu	Ala	Tyr	Ala	
			275					280					285				
5	act	atc	ttg	cat	tct	gcc	caa	ttt	tat	gtc	tgt	gga	gcc	att	gca	gct	912
	Thr	Ile	Leu	His	Ser	Ala	Gln	Phe	Tyr	Val	Cys	Gly	Ala	Ile	Ala	Ala	
		290					295					300					
10	gca	cag	agc	att	cga	atg	tca	ggc	tct	act	cgt	gat	ctg	gtc	ata	ctt	960
	Ala	Gln	Ser	Ile	Arg	Met	Ser	Gly	Ser	Thr	Arg	Asp	Leu	Val	Ile	Leu	
	305					310					315					320	
15	gtt	gat	gaa	acg	ata	agc	gaa	tac	cat	aaa	agt	ggc	ttg	gta	gct	gct	1008
	Val	Asp	Glu	Thr	Ile	Ser	Glu	Tyr	His	Lys	Ser	Gly	Leu	Val	Ala	Ala	
					325					330					335		
20	gga	tgg	aag	att	caa	atg	ttt	caa	aga	atc	agg	aac	ccg	aat	gct	gta	1056
	Gly	Trp	Lys	Ile	Gln	Met	Phe	Gln	Arg	Ile	Arg	Asn	Pro	Asn	Ala	Val	
				340					345					350			
	cca	aat	gcc	tac	aac	gaa	tgg	aac	tac	agc	aag	ttt	cgt	ctt	tgg	caa	1104
	Pro	Asn	Ala	Tyr	Asn	Glu	Trp	Asn	Tyr	Ser	Lys	Phe	Arg	Leu	Trp	Gln	
			355					360					365				
25	ctg	act	gaa	tac	agt	aag	atc	atc	ttc	atc	gat	gca	gac	atg	ctt	atc	1152
	Leu	Thr	Glu	Tyr	Ser	Lys	Ile	Ile	Phe	Ile	Asp	Ala	Asp	Met	Leu	Ile	
		370					375					380					
30	ctg	aga	aac	att	gat	ttc	ctc	ttc	gag	ttc	cct	gag	ata	tca	gca	act	1200
	Leu	Arg	Asn	Ile	Asp	Phe	Leu	Phe	Glu	Phe	Pro	Glu	Ile	Ser	Ala	Thr	
	385					390					395					400	
35	gga	aac	aat	gct	acg	ctc	ttc	aac	tct	ggc	cta	atg	gtg	gtt	gag	cca	1248
	Gly	Asn	Asn	Ala	Thr	Leu	Phe	Asn	Ser	Gly	Leu	Met	Val	Val	Glu	Pro	
				405						410					415		
40	tct	aat	tca	aca	ttc	cag	tta	cta	atg	gat	aac	att	aat	gaa	gtt	gtg	1296
	Ser	Asn	Ser	Thr	Phe	Gln	Leu	Leu	Met	Asp	Asn	Ile	Asn	Glu	Val	Val	
				420					425					430			
	tct	tac	aac	gga	gga	gac	caa	ggt	tac	ctt	aac	gag	ata	ttc	aca	tgg	1344
	Ser	Tyr	Asn	Gly	Gly	Asp	Gln	Gly	Tyr	Leu	Asn	Glu	Ile	Phe	Thr	Trp	
			435					440					445				
45	tgg	cat	cgg	att	cca	aaa	cac	atg	aat	ttc	ttg	aag	cat	ttc	tgg	gaa	1392
	Trp	His	Arg	Ile	Pro	Lys	His	Met	Asn	Phe	Leu	Lys	His	Phe	Trp	Glu	
		450					455					460					
50	gga	gac	gaa	cct	gag	att	aaa	aaa	atg	aag	acg	agt	cta	ttt	gga	gct	1440
	Gly	Asp	Glu	Pro	Glu	Ile	Lys	Lys	Met	Lys	Thr	Ser	Leu	Phe	Gly	Ala	
	465					470					475					480	
55	gat	cct	ccg	atc	cta	tac	gtt	ctt	cat	tac	cta	ggc	tat	aac	aaa	ccc	1488
	Asp	Pro	Pro	Ile	Leu	Tyr	Val	Leu	His	Tyr	Leu	Gly	Tyr	Asn	Lys	Pro	
					485					490					495		
60	tgg	tta	tgc	ttc	aga	gac	tat	gac	tgc	aat	tgg	aat	gtc	gat	att	ttc	1536
	Trp	Leu	Cys	Phe	Arg	Asp	Tyr	Asp	Cys	Asn	Trp	Asn	Val	Asp	Ile	Phe	
				500					505					510			
	cag	gaa	ttt	gct	agt	gac	gag	gct	cat	aaa	acc	tgg	tgg	aga	gtg	cac	1584

24

	Gln	Glu	Phe	Ala	Ser	Asp	Glu	Ala	His	Lys	Thr	Trp	Trp	Arg	Val	His	
			515					520					525				
5	gac	gca	atg	cct	gaa	aac	ttg	cat	aag	ttc	tgt	cta	cta	aga	tcg	aaa	1632
	Asp	Ala	Met	Pro	Glu	Asn	Leu	His	Lys	Phe	Cys	Leu	Leu	Arg	Ser	Lys	
		530					535					540					
10	cag	aag	gcg	caa	ctt	gaa	tgg	gat	agg	aga	caa	gca	gag	aaa	ggg	aac	1680
	Gln	Lys	Ala	Gln	Leu	Glu	Trp	Asp	Arg	Arg	Gln	Ala	Glu	Lys	Gly	Asn	
	545					550					555					560	
15	tac	aaa	gat	gga	cat	tgg	aag	ata	aag	atc	aaa	gac	aag	aga	ctt	aag	1728
	Tyr	Lys	Asp	Gly	His	Trp	Lys	Ile	Lys	Ile	Lys	Asp	Lys	Arg	Leu	Lys	
					565					570					575		
	act	tgt	ttc	gaa	gat	ttc	tgc	ttt	tgg	gag	agt	atg	ctt	tgg	cat	tgg	1776
	Thr	Cys	Phe	Glu	Asp	Phe	Cys	Phe	Trp	Glu	Ser	Met	Leu	Trp	His	Trp	
				580					585					590			
20	ggt	gag	acg	aac	tct	acc	aac	aat	tct	tcc	acc	acc	acc	act	tca	tca	1824
	Gly	Glu	Thr	Asn	Ser	Thr	Asn	Asn	Ser	Ser	Thr	Thr	Thr	Thr	Ser	Ser	
			595					600					605				
25	ccg	ccg	cat	aaa	acc	gct	ctc	cct	tcc	ctg	tga						1857
	Pro	Pro	His	Lys	Thr	Ala	Leu	Pro	Ser	Leu							
		610					615										
30	<210> 12																
	<211> 618																
	<212> PRT																
	<213> Arabidopsis thaliana																
35	<400> 12																
	Met	Ile	Pro	Ser	Ser	Ser	Pro	Met	Glu	Ser	Arg	His	Arg	Leu	Ser	Phe	
	1				5					10					15		
	Ser	Lys	Glu	Lys	Thr	Ser	Arg	Arg	Arg	Phe	Gln	Arg	Ile	Glu	Lys	Gly	
				20					25					30			
40	Val	Lys	Phe	Asn	Thr	Leu	Lys	Leu	Val	Leu	Ile	Cys	Ile	Met	Leu	Gly	
		35				40					45						
	Ala	Leu	Phe	Thr	Ile	Tyr	Arg	Phe	Arg	Tyr	Pro	Pro	Leu	Gln	Ile	Pro	
		50				55					60						
	Glu	Ile	Pro	Thr	Ser	Phe	Gly	Leu	Thr	Thr	Asp	Pro	Arg	Tyr	Val	Ala	
		65			70					75						80	
45	Thr	Ala	Glu	Ile	Asn	Trp	Asn	His	Met	Ser	Asn	Leu	Val	Glu	Lys	His	
				85					90						95		
	Val	Phe	Gly	Arg	Ser	Glu	Tyr	Gln	Gly	Ile	Gly	Leu	Ile	Asn	Leu	Asn	
			100					105					110				
50	Asp	Asn	Glu	Ile	Asp	Arg	Phe	Lys	Glu	Val	Thr	Lys	Ser	Asp	Cys	Asp	
		115						120					125				
	His	Val	Ala	Leu	His	Leu	Asp	Tyr	Ala	Ala	Lys	Asn	Ile	Thr	Trp	Glu	
		130				135					140						
	Ser	Leu	Tyr	Pro	Glu	Trp	Ile	Asp	Glu	Val	Glu	Glu	Phe	Glu	Val	Pro	
		145			150					155						160	
55	Thr	Cys	Pro	Ser	Leu	Pro	Leu	Ile	Gln	Ile	Pro	Gly	Lys	Pro	Arg	Ile	
				165				170							175		
	Asp	Leu	Val	Ile	Ala	Lys	Leu	Pro	Cys	Asp	Lys	Ser	Gly	Lys	Trp	Ser	
			180					185					190				
60	Arg	Asp	Val	Ala	Arg	Leu	His	Leu	Gln	Leu	Ala	Ala	Ala	Arg	Val	Ala	
		195					200					205					
	Ala	Ser	Ser	Lys	Gly	Leu	His	Asn	Val	His	Val	Ile	Leu	Val	Ser	Asp	

25

	210		215		220
	Cys Phe Pro Ile Pro Asn Leu Phe Thr Gly Gln Glu Leu Val Ala Arg				
	225		230		235
5	Gln Gly Asn Ile Trp Leu Tyr Lys Pro Asn Leu His Gln Leu Arg Gln				240
		245		250	
	Lys Leu Gln Leu Pro Val Gly Ser Cys Glu Leu Ser Val Pro Leu Gln				255
		260		265	
	Ala Lys Asp Asn Phe Tyr Ser Ala Gly Ala Lys Lys Glu Ala Tyr Ala				270
		275		280	
10	Thr Ile Leu His Ser Ala Gln Phe Tyr Val Cys Gly Ala Ile Ala Ala				285
		290		295	
	Ala Gln Ser Ile Arg Met Ser Gly Ser Thr Arg Asp Leu Val Ile Leu				300
		305		310	
	Val Asp Glu Thr Ile Ser Glu Tyr His Lys Ser Gly Leu Val Ala Ala				315
15			325		330
	Gly Trp Lys Ile Gln Met Phe Gln Arg Ile Arg Asn Pro Asn Ala Val				335
		340		345	
	Pro Asn Ala Tyr Asn Glu Trp Asn Tyr Ser Lys Phe Arg Leu Trp Gln				350
		355		360	
20	Leu Thr Glu Tyr Ser Lys Ile Ile Phe Ile Asp Ala Asp Met Leu Ile				365
		370		375	
	Leu Arg Asn Ile Asp Phe Leu Phe Glu Phe Pro Glu Ile Ser Ala Thr				380
		385		390	
	Gly Asn Asn Ala Thr Leu Phe Asn Ser Gly Leu Met Val Val Glu Pro				395
25			405		410
	Ser Asn Ser Thr Phe Gln Leu Leu Met Asp Asn Ile Asn Glu Val Val				415
		420		425	
	Ser Tyr Asn Gly Gly Asp Gln Gly Tyr Leu Asn Glu Ile Phe Thr Trp				430
		435		440	
30	Trp His Arg Ile Pro Lys His Met Asn Phe Leu Lys His Phe Trp Glu				445
		450		455	
	Gly Asp Glu Pro Glu Ile Lys Lys Met Lys Thr Ser Leu Phe Gly Ala				460
		465		470	
	Asp Pro Pro Ile Leu Tyr Val Leu His Tyr Leu Gly Tyr Asn Lys Pro				475
35			485		490
	Trp Leu Cys Phe Arg Asp Tyr Asp Cys Asn Trp Asn Val Asp Ile Phe				495
		500		505	
	Gln Glu Phe Ala Ser Asp Glu Ala His Lys Thr Trp Trp Arg Val His				510
		515		520	
40	Asp Ala Met Pro Glu Asn Leu His Lys Phe Cys Leu Leu Arg Ser Lys				525
		530		535	
	Gln Lys Ala Gln Leu Glu Trp Asp Arg Arg Gln Ala Glu Lys Gly Asn				540
		545		550	
	Tyr Lys Asp Gly His Trp Lys Ile Lys Ile Lys Asp Lys Arg Leu Lys				555
45			565		570
	Thr Cys Phe Glu Asp Phe Cys Phe Trp Glu Ser Met Leu Trp His Trp				575
		580		585	
	Gly Glu Thr Asn Ser Thr Asn Asn Ser Ser Thr Thr Thr Ser Ser				590
		595		600	
50	Pro Pro His Lys Thr Ala Leu Pro Ser Leu				605
		610		615	

55 <210> 13
 <211> 2130
 <212> DNA
 <213> Arabidopsis thaliana

60 <220>
 <221> CDS

<222> (1)..(1980)

<400> 13

5	atg gca aac tct ccc gct gct cct gca ccc acc acc aca acc ggt ggt	48
	Met Ala Asn Ser Pro Ala Ala Pro Ala Pro Thr Thr Thr Thr Gly Gly	
	1 5 10 15	
10	gac tcc cgg cga cgc ctc tcc gcg tcc ata gaa gca ata tgc aag agg	96
	Asp Ser Arg Arg Leu Ser Ala Ser Ile Glu Ala Ile Cys Lys Arg	
	20 25 30	
15	aga ttc cgg aga aat agc aaa gga ggt ggc aga tcg gat atg gtg aaa	144
	Arg Phe Arg Arg Asn Ser Lys Gly Gly Gly Arg Ser Asp Met Val Lys	
	35 40 45	
20	ccg ttt aat atc ata aat ttt tcg aca caa gac aaa aac agt agt tgt	192
	Pro Phe Asn Ile Ile Asn Phe Ser Thr Gln Asp Lys Asn Ser Ser Cys	
	50 55 60	
25	tgt tgt ttc acc aag ttt cag atc gtg aag ctt ctc ttg ttt atc ctt	240
	Cys Cys Phe Thr Lys Phe Gln Ile Val Lys Leu Leu Leu Phe Ile Leu	
	65 70 75 80	
30	ctc tct gcc act ctc ttc acc att atc tat tct cct gaa gct tat cat	288
	Leu Ser Ala Thr Leu Phe Thr Ile Ile Tyr Ser Pro Glu Ala Tyr His	
	85 90 95	
35	cat tct ctt tcc cac tca tct tct cga tgg ata tgg aga aga caa gat	336
	His Ser Leu Ser His Ser Ser Ser Arg Trp Ile Trp Arg Arg Gln Asp	
	100 105 110	
40	cca cgt tac ttc tcg gat ctg gat ata aac tgg gac gat gtg act aaa	384
	Pro Arg Tyr Phe Ser Asp Leu Asp Ile Asn Trp Asp Asp Val Thr Lys	
	115 120 125	
45	acc ctt gag aac atc gaa gaa ggc cgt acg atc ggt gtc ttg aat ttt	432
	Thr Leu Glu Asn Ile Glu Glu Gly Arg Thr Ile Gly Val Leu Asn Phe	
	130 135 140	
50	gat tcg aac gag atc caa cga tgg aga gaa gta tcc aag agc aag gac	480
	Asp Ser Asn Glu Ile Gln Arg Trp Arg Glu Val Ser Lys Ser Lys Asp	
	145 150 155 160	
55	aat ggg gat gaa gaa aaa gtt gtt gta ttg aat cta gat tac gca gac	528
	Asn Gly Asp Glu Glu Lys Val Val Val Leu Asn Leu Asp Tyr Ala Asp	
	165 170 175	
60	aag aat gtg act tgg gac gca cta tat cca gag tgg atc gat gag gag	576
	Lys Asn Val Thr Trp Asp Ala Leu Tyr Pro Glu Trp Ile Asp Glu Glu	
	180 185 190	
65	caa gaa aca gag gtc cct gtt tgt cct aat atc ccg aac att aag gta	624
	Gln Glu Thr Glu Val Pro Val Cys Pro Asn Ile Pro Asn Ile Lys Val	
	195 200 205	
70	cct aca aga aga ctc gat ctg atc gtc gtg aaa ctt cct tgt cgg aaa	672
	Pro Thr Arg Arg Leu Asp Leu Ile Val Val Lys Leu Pro Cys Arg Lys	
	210 215 220	
75	gaa ggg aat tgg tcg aga gac gtc ggg aga ttg cat cta cag cta gcg	720
	Glu Gly Asn Trp Ser Arg Asp Val Gly Arg Leu His Leu Gln Leu Ala	

27

	225	230	235	240	
5	gct gca act gtg gcg gct tcg gcc aaa ggg ttt ttc agg gga cat gtg Ala Ala Thr Val Ala Ala Ser Ala Lys Gly Phe Phe Arg Gly His Val	245	250	255	768
10	ttt ttt gta tct aga tgc ttt ccg att ccg aat ctt ttc cgg tgt aaa Phe Phe Val Ser Arg Cys Phe Pro Ile Pro Asn Leu Phe Arg Cys Lys	260	265	270	816
15	gat ctt gtg tct cgg aga ggc gat gtt tgg ttg tac aaa cct aat ctt Asp Leu Val Ser Arg Arg Gly Asp Val Trp Leu Tyr Lys Pro Asn Leu	275	280	285	864
20	gat acc ttg aga gac aag ctt cag ctg cct gta ggg tct tgt gag cta Asp Thr Leu Arg Asp Lys Leu Gln Leu Pro Val Gly Ser Cys Glu Leu	290	295	300	912
25	tct ctt cct ctt ggc atc caa gat agg cca agc tta gga aac cct aaa Ser Leu Pro Leu Gly Ile Gln Asp Arg Pro Ser Leu Gly Asn Pro Lys	305	310	315	960
30	aga gaa gct tac gca aca att ctt cac tca gct cac gtt tac gtc tgc Arg Glu Ala Tyr Ala Thr Ile Leu His Ser Ala His Val Tyr Val Cys	325	330	335	1008
35	ggg gca atc gcc gcg gct cag agc ata aga cag tct ggt tgc acg aga Gly Ala Ile Ala Ala Ala Gln Ser Ile Arg Gln Ser Gly Ser Thr Arg	340	345	350	1056
40	gac ctt gtt atc ctt gtt gat gac aac atc agc ggt tac cac cgg agt Asp Leu Val Ile Leu Val Asp Asp Asn Ile Ser Gly Tyr His Arg Ser	355	360	365	1104
45	gga cta gaa gcc gcg ggt tgg caa atc cgg acg ata cag agg att cga Gly Leu Glu Ala Ala Gly Trp Gln Ile Arg Thr Ile Gln Arg Ile Arg	370	375	380	1152
50	aac cct aag gca gag aaa gat gct tac aac gaa tgg aac tac agc aag Asn Pro Lys Ala Glu Lys Asp Ala Tyr Asn Glu Trp Asn Tyr Ser Lys	385	390	395	1200
55	ttc cgg cta tgg cag ctg act gat tac gac aaa atc att ttc atc gac Phe Arg Leu Trp Gln Leu Thr Asp Tyr Asp Lys Ile Ile Phe Ile Asp	405	410	415	1248
60	gcg gat ctc tta atc ttg aga aac atc gat ttc ttg ttc tgc atg cct Ala Asp Leu Leu Ile Leu Arg Asn Ile Asp Phe Leu Phe Ser Met Pro	420	425	430	1296
65	gag atc tca gct aca gga aac aat gga act ctg ttt aat tca gga gtt Glu Ile Ser Ala Thr Gly Asn Asn Gly Thr Leu Phe Asn Ser Gly Val	435	440	445	1344
70	atg gtg atc gag cct tgc aac tgt acg ttt cag ctt ctg atg gaa cat Met Val Ile Glu Pro Cys Asn Cys Thr Phe Gln Leu Leu Met Glu His	450	455	460	1392
75	ata aac gag att gag tct tat aac ggt gga gat caa ggt tac tta aac Ile Asn Glu Ile Glu Ser Tyr Asn Gly Gly Asp Gln Gly Tyr Leu Asn	465	470	475	1440

	gag gta ttc aca tgg tgg cac cgg att cca aaa cat atg aat ttc ttg	1488
	Glu Val Phe Thr Trp Trp His Arg Ile Pro Lys His Met Asn Phe Leu	
	485 490 495	
5	aag cat ttt tgg att ggc gat gaa gat gac gcg aaa cgc aag aaa aca	1536
	Lys His Phe Trp Ile Gly Asp Glu Asp Asp Ala Lys Arg Lys Lys Thr	
	500 505 510	
10	gag ctt ttt gga gca gag cct cct gtt ctt tat gtt ctt cat tac ctt	1584
	Glu Leu Phe Gly Ala Glu Pro Pro Val Leu Tyr Val Leu His Tyr Leu	
	515 520 525	
15	ggg atg aag ccg tgg tta tgt tac cgt gac tac gac tgt aac ttc aac	1632
	Gly Met Lys Pro Trp Leu Cys Tyr Arg Asp Tyr Asp Cys Asn Phe Asn	
	530 535 540	
20	tcc gac ata ttc gtt gag ttt gct acc gat atc gct cat cga aaa tgg	1680
	Ser Asp Ile Phe Val Glu Phe Ala Thr Asp Ile Ala His Arg Lys Trp	
	545 550 555 560	
	tgg atg gtc cac gac gcc atg cca cag gaa ctt cac caa ttc tgt tac	1728
	Trp Met Val His Asp Ala Met Pro Gln Glu Leu His Gln Phe Cys Tyr	
	565 570 575	
25	ttg cga tcc aag caa aag gca cag ctg gaa tat gat cgc cgg caa gca	1776
	Leu Arg Ser Lys Gln Lys Ala Gln Leu Glu Tyr Asp Arg Arg Gln Ala	
	580 585 590	
30	gag gcc gca aat tat gcc gac ggt cat tgg aaa ata aga gta aag gac	1824
	Glu Ala Ala Asn Tyr Ala Asp Gly His Trp Lys Ile Arg Val Lys Asp	
	595 600 605	
35	ccg aga ttc aaa att tgc atc gac aaa tta tgt aat tgg aaa agt atg	1872
	Pro Arg Phe Lys Ile Cys Ile Asp Lys Leu Cys Asn Trp Lys Ser Met	
	610 615 620	
40	ctg cgg cat tgg ggc gaa tca aat tgg act gac tac gag tct ttt gtt	1920
	Leu Arg His Trp Gly Glu Ser Asn Trp Thr Asp Tyr Glu Ser Phe Val	
	625 630 635 640	
	ccc acc cca cca gcc att acc gta gac cgg aga tca tca ctt ccc ggc	1968
	Pro Thr Pro Pro Ala Ile Thr Val Asp Arg Arg Ser Ser Leu Pro Gly	
	645 650 655	
45	cat aac ttg tga cgcaataatt atacatactt attaatggat ttcattgagtt	2020
	His Asn Leu	
	660	
50	tttttggtttg aattgttgct gcgagattag gtgaatatca gttgtgtaac tataatctttt	2080
	tcctatagtt tgttcaaatt gaataaaaca tttttttgca gtttaaccac	2130
55	<210> 14	
	<211> 659	
	<212> PRT	
	<213> Arabidopsis thaliana	
60	<400> 14	
	Met Ala Asn Ser Pro Ala Ala Pro Ala Pro Thr Thr Thr Thr Gly Gly	

29

		1			5				10				15				
		Asp	Ser	Arg	Arg	Arg	Leu	Ser	Ala	Ser	Ile	Glu	Ala	Ile	Cys	Lys	Arg
					20				25					30			
5		Arg	Phe	Arg	Arg	Asn	Ser	Lys	Gly	Gly	Gly	Arg	Ser	Asp	Met	Val	Lys
			35						40					45			
		Pro	Phe	Asn	Ile	Ile	Asn	Phe	Ser	Thr	Gln	Asp	Lys	Asn	Ser	Ser	Cys
			50				55						60				
		Cys	Cys	Phe	Thr	Lys	Phe	Gln	Ile	Val	Lys	Leu	Leu	Leu	Phe	Ile	Leu
			65			70					75						80
10		Leu	Ser	Ala	Thr	Leu	Phe	Thr	Ile	Ile	Tyr	Ser	Pro	Glu	Ala	Tyr	His
					85						90					95	
		His	Ser	Leu	Ser	His	Ser	Ser	Ser	Arg	Trp	Ile	Trp	Arg	Arg	Gln	Asp
				100					105					110			
		Pro	Arg	Tyr	Phe	Ser	Asp	Leu	Asp	Ile	Asn	Trp	Asp	Asp	Val	Thr	Lys
15				115					120					125			
		Thr	Leu	Glu	Asn	Ile	Glu	Glu	Gly	Arg	Thr	Ile	Gly	Val	Leu	Asn	Phe
			130				135						140				
		Asp	Ser	Asn	Glu	Ile	Gln	Arg	Trp	Arg	Glu	Val	Ser	Lys	Ser	Lys	Asp
					145		150					155					160
20		Asn	Gly	Asp	Glu	Glu	Lys	Val	Val	Val	Leu	Asn	Leu	Asp	Tyr	Ala	Asp
					165						170					175	
		Lys	Asn	Val	Thr	Trp	Asp	Ala	Leu	Tyr	Pro	Glu	Trp	Ile	Asp	Glu	Glu
				180						185					190		
		Gln	Glu	Thr	Glu	Val	Pro	Val	Cys	Pro	Asn	Ile	Pro	Asn	Ile	Lys	Val
25				195					200					205			
		Pro	Thr	Arg	Arg	Leu	Asp	Leu	Ile	Val	Val	Lys	Leu	Pro	Cys	Arg	Lys
			210				215						220				
		Glu	Gly	Asn	Trp	Ser	Arg	Asp	Val	Gly	Arg	Leu	His	Leu	Gln	Leu	Ala
			225			230						235					240
30		Ala	Ala	Thr	Val	Ala	Ala	Ser	Ala	Lys	Gly	Phe	Phe	Arg	Gly	His	Val
					245						250					255	
		Phe	Phe	Val	Ser	Arg	Cys	Phe	Pro	Ile	Pro	Asn	Leu	Phe	Arg	Cys	Lys
				260						265					270		
		Asp	Leu	Val	Ser	Arg	Arg	Gly	Asp	Val	Trp	Leu	Tyr	Lys	Pro	Asn	Leu
35				275					280					285			
		Asp	Thr	Leu	Arg	Asp	Lys	Leu	Gln	Leu	Pro	Val	Gly	Ser	Cys	Glu	Leu
			290				295						300				
		Ser	Leu	Pro	Leu	Gly	Ile	Gln	Asp	Arg	Pro	Ser	Leu	Gly	Asn	Pro	Lys
			305			310						315				320	
40		Arg	Glu	Ala	Tyr	Ala	Thr	Ile	Leu	His	Ser	Ala	His	Val	Tyr	Val	Cys
					325						330					335	
		Gly	Ala	Ile	Ala	Ala	Ala	Gln	Ser	Ile	Arg	Gln	Ser	Gly	Ser	Thr	Arg
				340						345					350		
		Asp	Leu	Val	Ile	Leu	Val	Asp	Asp	Asn	Ile	Ser	Gly	Tyr	His	Arg	Ser
45				355				360						365			
		Gly	Leu	Glu	Ala	Ala	Gly	Trp	Gln	Ile	Arg	Thr	Ile	Gln	Arg	Ile	Arg
			370				375						380				
		Asn	Pro	Lys	Ala	Glu	Lys	Asp	Ala	Tyr	Asn	Glu	Trp	Asn	Tyr	Ser	Lys
			385			390						395				400	
50		Phe	Arg	Leu	Trp	Gln	Leu	Thr	Asp	Tyr	Asp	Lys	Ile	Ile	Phe	Ile	Asp
					405						410					415	
		Ala	Asp	Leu	Leu	Ile	Leu	Arg	Asn	Ile	Asp	Phe	Leu	Phe	Ser	Met	Pro
				420					425						430		
		Glu	Ile	Ser	Ala	Thr	Gly	Asn	Asn	Gly	Thr	Leu	Phe	Asn	Ser	Gly	Val
55				435					440					445			
		Met	Val	Ile	Glu	Pro	Cys	Asn	Cys	Thr	Phe	Gln	Leu	Leu	Met	Glu	His
			450				455						460				
		Ile	Asn	Glu	Ile	Glu	Ser	Tyr	Asn	Gly	Gly	Asp	Gln	Gly	Tyr	Leu	Asn
			465			470						475				480	
60		Glu	Val	Phe	Thr	Trp	Trp	His	Arg	Ile	Pro	Lys	His	Met	Asn	Phe	Leu
					485						490					495	

30

Lys His Phe Trp Ile Gly Asp Glu Asp Asp Ala Lys Arg Lys Lys Thr
 500 505 510
 Glu Leu Phe Gly Ala Glu Pro Pro Val Leu Tyr Val Leu His Tyr Leu
 515 520 525
 5 Gly Met Lys Pro Trp Leu Cys Tyr Arg Asp Tyr Asp Cys Asn Phe Asn
 530 535 540
 Ser Asp Ile Phe Val Glu Phe Ala Thr Asp Ile Ala His Arg Lys Trp
 545 550 555 560
 10 Trp Met Val His Asp Ala Met Pro Gln Glu Leu His Gln Phe Cys Tyr
 565 570 575
 Leu Arg Ser Lys Gln Lys Ala Gln Leu Glu Tyr Asp Arg Arg Gln Ala
 580 585 590
 Glu Ala Ala Asn Tyr Ala Asp Gly His Trp Lys Ile Arg Val Lys Asp
 595 600 605
 15 Pro Arg Phe Lys Ile Cys Ile Asp Lys Leu Cys Asn Trp Lys Ser Met
 610 615 620
 Leu Arg His Trp Gly Glu Ser Asn Trp Thr Asp Tyr Glu Ser Phe Val
 625 630 635 640
 20 Pro Thr Pro Pro Ala Ile Thr Val Asp Arg Arg Ser Ser Leu Pro Gly
 645 650 655
 His Asn Leu

25 <210> 15
 <211> 15294
 <212> DNA
 <213> Artificial Sequence

30 <220>
 <220>
 <223> Description of Artificial Sequence: vector

35 <400> 15
 ggccggggagg gttcgagaag ggggggcacc ccccttcggc gtgcgcggtc acgcgcacag 60
 ggcgcagccc tggttaaaaa caaggtttat aaatatgggt ttaaaagcag gttaaaagac 120
 aggttagcgg tggccgaaaa acgggcgga acccttgcaa atgctggatt ttctgcctgt 180
 ggacagcccc tcaaattgtca ataggtgcgc ccctcatctg tcagcactct gcccctcaag 240
 40 tgtcaaggat cgcgcccctc atctgtcagt agtcgcgccc ctcaagtgtc aataccgcag 300
 ggcacttatc cccaggcttg tccacatcat ctgtgggaaa ctgcgtaaa atcaggcggt 360
 ttgcgcgatt tgcgaggctg gccagctcca cgtcgccggc cgaaatcgag cctgcccctc 420
 atctgtcaac gccgcgcgg gtgagtcggc cctcaagtg tcaacgtccg cccctcatct 480
 gtcatgtagg gccaaagttt ccgcgaggt tccacaacgc cggcgccgc ggtgtctcgc 540
 45 acacgggttc gacggcggtt ctggcgcggt tgcagggcca tagacggccg ccagcccagc 600
 ggcgagggca accagcccgg tgagcgtcgc aaaggcgctc ggtcttgctt tgctcgtcgg 660
 tgatgtactt caccagctcc gcgaagtgcg tcttcttgat ggagcgcagt gggacgtgct 720
 tggcaatcac gcgcaccccc cggccgtttt agcgggctaaa aaagtcatgg ctctgccctc 780
 gggcggaacca cgcccatcat gaccttgcca agctcgtcct gcttctcttc gatcttcgcc 840
 50 agcagggcga ggatcgtggc atcacccaac cgcgcgcgtg gcgggtcgtc ggtgagccag 900
 agtttccagca ggccgcccag gcggcccagg tcgccattga tgcgggccag ctgcgcggacg 960
 tgcctcatagt ccacgacgcc cgtgattttg tagccctggc cgacggccag caggtaggcc 1020
 gacaggctca tgccggccgc cgcgcctttt tcctcaatcg ctcttcgttc gtctggaagg 1080
 cagtacacct tgataggtgg gctgcccttc ctggttggct tggtttcacg agccatccgc 1140
 55 ttgccctcat ctgttacgcc ggcggtagcc ggccagcctc gcagagcagg attcccgttg 1200
 agcaccgcca ggtgcgaata agggacagt aagaaggaa acccgctcgc ggggtggcct 1260
 acttcacctt tcctgcccgg ctgacgcctg tggatacacc aaggaaagtc tacacgaacc 1320
 ctttggaata atcctgtata tcgtgcgaaa aaggatggat ataccgaaaa aatcgctata 1380
 atgaccccga agcagggtta tgcagcgga aagcgccacg cttcccgaag ggagaaaggc 1440
 60 ggacaggtat ccggtaaagc gcagggtcgg aacaggagag cgcacgaggg agcttcagg 1500
 gggaaacgcc tggatatctt atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg 1560

	atttttgtga	tgctcgtcag	gggggaggag	cctatggaaa	aacgccagca	acgcggcctt	1620
	tttacgggtc	ctggcctttt	gctggccttt	tgctcacatg	ttctttcctg	cgttatcccc	1680
	tgattctgtg	gataaccgta	ttaccgcctt	tgagttagct	gataccgctc	gccgcagccg	1740
5	aacgaccgag	cgcagcgagt	cagttagcga	ggaagcggaa	gagcgccaga	aggccgcccag	1800
	agaggccgag	cgcggccgtg	aggcttgga	gctagggcag	ggcatgaaaa	agcccgtagc	1860
	gggctgctac	gggcgtctga	cgcgggtgaa	agggggaggg	gatgttgtct	acatggctct	1920
	gctgtagtga	gtgggttgcg	ctccggtcagc	ggctctgatc	aatcgtcacc	ctttctcggt	1980
	ccttcaacgt	tcctgacaac	gagcctcctt	ttcgccaatc	catcgacaat	caccgcgagt	2040
	ccctgctcga	acgctgcgtc	cggaccggct	tcgtcgaagg	cgtctatcgc	ggcccgcac	2100
10	agcggcgaga	gcgagcctg	ttcaacgggt	cgcgcgcgt	cgcggccatc	gctgtcgccg	2160
	gcctgctcct	caagcacggc	cccaacagt	aagtagctga	ttgtcatcag	cgcattgacg	2220
	gcgtccccgg	ccgaaaaacc	cgcctcgcag	aggaagcgaa	gctgcgcgtc	ggcggtttcc	2280
	atctgcggtg	cgcgcggctg	cgtgccggca	tggatgcgcg	cgcctatcgc	gtaggcgagc	2340
	agcgctgcc	tgaagctgcg	ggcattcccg	atcagaaatg	agcgccagtc	gtcgtcggt	2400
15	ctcggcaccg	aatgcgtatg	attctccgcc	agcatggctt	cggccagtgc	gtcgagcagc	2460
	gcccgtttgt	tcctgaagt	ccagtaaagc	gcccgtgct	gaaccccaa	ccgttccgcc	2520
	gctttgcgtg	tcgtcagacc	gtctacgcgc	acctcggtca	acagggtccag	ggcgccagcg	2580
	atcactgtat	tcggctgcaa	ctttgtcatg	cttgacactt	tatcactgat	aaacataata	2640
20	tgtccacca	cttatcagtg	ataaagaatc	cgcgcgttca	atcggaccag	cggaggctgg	2700
	tcgggaggcc	agacgtgaaa	cccaacatac	ccctgatcgt	aattctgagc	actgtcgccg	2760
	tcgacgctgt	cggcatcggc	ctgattatgc	cgggtgctgcc	gggcctcctg	cgcgatctgg	2820
	ttcactcgaa	cgcagtcacc	gcccactatg	gcattctgct	ggcgctgtat	gcggttggtg	2880
	aatttgcttg	cgcacctgtg	ctgggcgcgc	tgctcgatcg	tttcggggcg	cggccaatct	2940
	tgctcgtctc	gctggccggc	gccaagatct	ggggaaccct	gtggttgga	tgcacataca	3000
25	aatggacgaa	cggataaaac	ttttcacgcc	cttttaata	tccgattatt	ctcaataaacg	3060
	ctcttttctc	ttagggtttac	ccgccaatat	atcctgtcaa	acactgatag	tttaaaactga	3120
	aggcgggaaa	cgacaatctg	atcatgagcg	gagaattaag	ggagtcacgt	tatgaccccc	3180
	gccgatgacg	cgggacaagc	cgtttttacgt	ttggaactga	cagaaccgca	acgttgaagg	3240
	agccactcag	ccgatctgaa	ttcccgatct	agtaacatag	atgacaccgc	gcgcgataat	3300
30	ttatcctagt	ttgcgcgcta	tattttgttt	tctatcgct	attaaatgta	taattgcggg	3360
	actctaata	taaaaaccca	tctcataaat	aacgtcatgc	attacatgtt	aattattaca	3420
	tgcttaacgt	aattcaacag	aaattatatg	ataatcatcg	caagaccggc	aacaggattc	3480
	aactttaaga	aaactttattg	ccaaatgttt	gaacgatcgg	ggaaattcga	gctcggtagc	3540
	atcatgttac	aaactttttt	gctgtgagca	gtagatattg	aaaccgggag	gacctaaggt	3600
35	atctgataga	taaagcacat	agcttggtt	tacaggttct	ggtggatgta	gttcacagtc	3660
	atgcaagcaa	taatgccact	gatggcctca	atggctttga	tattggccaa	ggttctcaag	3720
	aatcctactt	tcagtctgga	gagcaagggt	accataagtt	gtgggatagc	aggctgttca	3780
	actatgccaa	ttgggaggtt	cttcgtttcc	ttctttccaa	cttgaggtgg	tggctagaag	3840
	agtataaactt	tgacggattt	cgatttgatg	gaataacttc	tatgctgtat	gttcacatcg	3900
40	gaatcaatat	gggattttaca	ggaaactata	atgagtattt	cagcgaggct	acagatgttg	3960
	atgctgtggt	ctattttaatg	ttggccaata	atctgatcca	caagattttc	ccagacgcaa	4020
	ctgttatggc	cgaagatgtt	tctggtatgc	cgggccttgg	ccggcctgtt	tctgagggag	4080
	gaattggttt	tgattaccgc	ctggcaatgg	caatcccaga	taagtggata	gattatttaa	4140
	agaataaaaa	tgatgaagat	tggtccatga	aggaagtaac	atcgagtttg	acaaatagga	4200
45	gatatacaga	gaagtgtata	gcataatgcg	agagccatga	tcagtctatt	gtcgggtgaca	4260
	agaccattgc	atttctccta	atggacaaag	agatgtattc	tggcatgtct	tgcttgacag	4320
	atgcttctcc	tgattattgat	cgaggaattg	cgcttcacaa	gatgatccat	tttttcacaa	4380
	tggccttggg	aggagagggg	tacctcaatt	tcatgggtaa	cgagtttggc	catcctgagt	4440
	ggattgactt	ccctagagag	ggcaataatt	ggtgttatga	caaatgtaga	cgccagtggg	4500
50	accttgccga	tagcgaacac	ttgagataca	agtttatgaa	tgcatattgat	agagctatga	4560
	attcgctcga	tgaaaagtcc	tcattcctcg	catcaggaaa	acagatagta	agcagcatgg	4620
	atgatgagaa	gaagggttgtt	gtggttgaa	gtggtgacct	ggtatttgta	ttcaacttcc	4680
	acccaaataa	cacatacgaa	gggtataaag	ttggatgtga	cttgccaggg	aagtacagag	4740
	ttgcaactgga	cagtgtatgt	tgggaatttg	gtggccatgg	aagagctggt	catgatgttg	4800
55	accatttcac	atcaccagaa	ggaataacctg	gagttccaga	aacaaatttc	aatggctcgtc	4860
	caaattcctt	caaagtgcgt	tctcctgcgc	gaacatgtgt	ggcttattac	agagttgacg	4920
	aacgcagtgc	agaaactgaa	gtttaccaga	cagacatttc	tagtgagcta	ctaccaacag	4980
	ccaatatoga	ggagagtgc	gagaaactta	aagattcgtt	atctacaaat	atcagtaaac	5040
	ttgacgaact	catgtcagaa	actgaagttt	accagacaga	catttctagt	gagctactac	5100
60	caacagccag	tatcgaggag	agtgcagaga	aacttaagaa	ttcattatct	acaaatatca	5160
	gtacgtgggt	atcattggat	gtgggattcc	cgcctcttta	attatggaaa	ctgggaggta	5220

	cttaggtatc	ttctctcaaa	tgcgagatgg	tggttggatg	agttcaaatt	tgatggattt	5280
	agattcgatg	gtgtgacatc	aataatgtat	actcaccacg	gattatcggg	gggattcact	5340
	gggaactaca	aggaatactt	tggactcgca	actgatgtgg	atgctgttgt	gtatctgatg	5400
	ctgggtcaacg	atcttattca	tgggcttttc	cagatgcaat	taccattggg	gaagatgtta	5460
5	gcggaatgcc	gacattttgt	attcccgttc	aagatggggg	tggttggctt	gactatcggc	5520
	tgcatatggc	aattgctgat	aaatggattg	agttgctcaa	gaaacgggat	gaggattgga	5580
	gagtgggtga	tattgttcat	acactgacaa	atagaagatg	gtcggaaaag	tgtgtttcat	5640
	acgctgaaag	tcatgatcaa	gctctagtctg	gtgataaaac	tatagcattc	tggtctgatg	5700
	acaaggatat	gtatgatttt	atggctctgg	atagaccatc	aacatcatta	atagatcgtg	5760
10	ggatagcatt	gcacaagatg	attaggcttg	taactatggg	attaggagga	gaaggggtacc	5820
	taaattttcat	gggaaatgaa	ttcgccacc	ctgagtggat	tgattttccct	agggctgaac	5880
	aacacctctc	tgatggctca	gtaattcccg	gaaaccaatt	cagttatgat	aatgcagac	5940
	ggagatttga	cctgggagat	gcagaatatt	taagataccg	tggttgcaa	gaatttgacc	6000
	gggctatgca	gtatcttgaa	gataaatatg	agtttatgac	ttcagaacac	cagttcatat	6060
15	cacgaaagga	tgaaggagat	aggatgattg	tatttgaaaa	aggaaaccta	gtttttgtct	6120
	ttaattttca	ctggacaaaa	agctattcag	actatcgcat	aggctgcctg	aagcctggaa	6180
	aatacaaggt	tccttggac	tcagatgatc	cactttttgg	tggtctcggg	agaattgatc	6240
	ataatgccga	atatttcacc	tttgaaggat	ggtatgtga	tcgtcctcgt	tcaattatgg	6300
	tgtatgcacc	tagtagaaca	gcagtgggtc	atgcactagt	agacaaagaa	gaagaagaag	6360
20	aagaagaagt	agcagtagta	gaagaagtag	tagtagaaga	agaatgaacg	aacttgtgat	6420
	cgcgttgaaa	gatttgaagg	ctacatagct	ctagagtcga	cctgcatgaa	atcagaaata	6480
	attggaggag	atgagtaaaa	gttaccactt	gttgagctgt	gtgagtgaat	gagtggagaat	6540
	gaggagggtc	ctgccttatt	tgtagcaggt	ttcagtgaac	cgtgtcaaga	gaatagcggg	6600
	tggttatccc	ttagcagaag	gcaactgtgg	acactgtatt	atagggaaat	gctcatcgac	6660
25	agtattatgg	gcctctctct	tgttgattca	cggctggact	tcaacttggg	ccttgcaatg	6720
	ggcccgctcg	gttctgtctc	ctagtatcta	aaaaactaaa	ccaactccct	ctcaccgcta	6780
	ccacttgaca	ttcctatgtc	tcgtgttaat	taaattatta	ttatagtaat	taaaaataat	6840
	atctaggtac	tggtactggg	ccctccctcc	actagaatat	tagttacttc	ccccttagct	6900
	ttgtattcca	aattactgta	aatatatttt	ctaatttttt	acgacaaaca	agatctaatt	6960
30	atgaatgcac	aattctaaag	gttgaataca	ttactttact	tggttttagc	tatattaagt	7020
	tgcattttag	tattaagatt	gagatgcatt	gttctattac	aaaattgata	cactgctaaa	7080
	ggaaggatgg	ttaaaaacaa	cattcaatgt	ttgttacatt	tcttcctatt	gtattttttt	7140
	tttaacgagc	ttcccgtata	catcataaca	tgtctccgtt	ccacttggca	ggaaaaaaa	7200
	atacccaaac	aggaagatac	tgtcaagtat	atccatagat	gaggacttaa	tggataggct	7260
35	tttcgaggat	tcataaatca	taatatctgg	cggaggagtc	aattaaatac	ttgtggtttg	7320
	tatcctgatt	actccgtcaa	cagccaaata	gaaaagtttg	aaaagagaga	aaggatttgg	7380
	tacaagatac	tggtgcattt	gttaagtaat	gaacaaaacg	gagtaacata	atcttctatc	7440
	tcgttaaagc	ttcacgctgc	cgcaagcact	cagggcgcaa	gggctgctaa	ggaagcggaa	7500
	cacgtagaaa	gccagtcgct	agaaaacggtg	ctgaccccg	atgaatgtca	gctactgggc	7560
40	tatctggaca	agggaaaacg	caagcgcaaa	gagaaagcag	gtagcttgca	gtgggcttac	7620
	atggcgatag	ctagactggg	cggttttatg	gacagcaagc	gaaccggaat	tgccagctgg	7680
	ggcgccctct	ggtaagggtg	ggaagccctg	caaagtaaac	tggatggctt	tcttgccgcc	7740
	aaggactctga	tggcgaggag	gtacaagatc	atgagcggag	aattaaggga	gtcagcttat	7800
	gacccccgct	gatgacgctg	gacaagccgt	tttacgtttg	gaactgacag	aaccgcaacg	7860
45	ttgaaggagc	cactcagccg	cgggtttctg	gagtttaatt	agctaagcac	atacgtcaga	7920
	aaccattatt	gcgcgttcaa	aagtcgccta	aggtcactat	cagctagcaa	atatttcttg	7980
	tcaaaaatgc	tccactgacg	ttccataaat	tcccctcggt	atccaattag	agtctcatat	8040
	tactctcaa	tccagatctc	gactctagtc	gagggcccat	gggagcttgg	attgaacaag	8100
	atggattgca	cgcaggttct	cggccgctt	gggtggagag	gctattcggc	tatgactggg	8160
50	cacaacagac	aatcggctgc	tctgatgccg	ccgtgttccg	gctgtcagcg	cagggcgcc	8220
	cggttctttt	tgtcaagacc	gacctgtccg	gtgccctgaa	tgaactgcag	gacgaggcag	8280
	cgggctatc	gtggctggcc	acgacgggct	ttccttgccg	agctgtgctc	gacgttgtca	8340
	ctgaagcggg	aagggaactg	ctgctatttg	gcgaagtgc	ggggcaggat	ctcctgtcat	8400
	ctcaccttgc	tcctgccgag	aaagtatcca	tcatggctga	tgcaatgcgg	cggctgcata	8460
55	cgttgatcc	ggctacctgc	ccattcgacc	accaagcgaa	acatcgcatc	gagcgagcac	8520
	gtactcggat	ggaagccggt	cttgctgatc	aggatgatct	ggacgaagag	catcaggggc	8580
	tcgcccagc	cgaactgttc	gccaggctca	aggcgccgat	gcccgcagcg	gaggatctcg	8640
	tcgtgaccca	tgggcgatgcc	tgcttgccga	atatcatggt	ggaaaatggc	cgtttttctg	8700
	gattcatcga	ctgtggccgg	ctgggtgtgg	cggaccgcta	tcaggacata	gcgttggtca	8760
60	cccgtgatat	tgctgaagag	cttggcgggc	aatggctga	ccgcttctc	gtgctttacg	8820
	gtatcgccgc	tcccgattcg	cagcgcatcg	ccttctatcg	ccttcttgac	gagttcttct	8880

	gagcgggacc	caagctagct	tcgacggatc	ccccgatgag	ctaagctagc	tatatcatca	8940
	atztatgtat	tacacataat	atcgcaactca	gtctttcatc	tacggcaatg	taccagctga	9000
	tataatcagt	tattgaaata	tttctgaatt	taaacttgca	tcaataaatt	tatgtttttg	9060
5	cttggactat	aatacctgac	ttgttatttt	atcaataaat	atttaaacta	tattttctttc	9120
	aagatgggaa	ttaattcact	ggccgtcgtt	ttacaacgtc	gtgactggga	aaaccctggc	9180
	gttaccacaac	ttaatcgct	tgcagcacat	ccccctttcg	ccagctggcg	taatagcgaa	9240
	gagggccgca	ccgatcgccc	ttcccaacag	ttgcgcagcc	tgaatggcgc	ccgctccttt	9300
	cgctttcttc	ccttcctttc	tcgccacgtt	cgccggcttt	ccccgtcaag	ctctaaatcg	9360
	ggggctccct	ttagggttcc	gatttagtgc	tttacggcac	ctcgacccca	aaaaacttga	9420
10	tttgggtgat	ggttcacgta	gtgggccatc	gccctgatag	acggtttttc	gccctttgac	9480
	gttggagtcc	acgttcttta	atagtggact	cttgttccaa	actggaacaa	cactcaacct	9540
	tatctcgggc	tattcttttg	atattataagg	gatttttgccg	atttcggaac	caccatcaaa	9600
	caggattttc	gcctgctggg	gcaaaccagc	gtggaccgct	tgctgcaact	ctctcagggc	9660
	caggcggtga	agggcaatca	gctgttgccc	gtctcactgg	tgaaaagaaa	aaccacccca	9720
15	gtacattaaa	aacgtccgca	atgtgttatt	aagttgtcta	agcgtcaatt	tgtttacacc	9780
	acaatatatc	ctgccaccag	ccagccaaca	gtccccgac	cggcagctcg	gcacaaaatc	9840
	accactcgat	acaggcagcc	catcagtcgc	ggacggcgtc	agcgggagag	ccgttgtaag	9900
	gcggcagact	ttgctcatgt	taccgatgct	attcgggaaga	acggcaacta	agctgcgggg	9960
	tttgaaacac	ggatgatctc	gcggagggtg	gcattgtgat	tgtaacgatg	acagagcggt	
20	10020						
	gctgcctgtg	atcaaataat	atctccctcg	cagagatccg	aattatcagc	cttcttattc	
	10080						
	atctctcgct	taaccgtgac	aggctgtcga	tcttgagaac	tatgccgaca	taataggaaa	
	10140						
25	tcgctggata	aagccgctga	ggaagctgag	tggcgctatt	tctttagaag	tgaacgttga	
	10200						
	cgatatcaac	tcccctatcc	attgctcacc	gaatgggtaca	ggtcggggac	ccgaagttcc	
	10260						
	gactgtcggc	ctgatgcac	cccggctgat	cgaccccgag	tctagatctg	gggctgagaa	
30	10320						
	agcccagtaa	ggaaacaact	gtaggttcga	gtcgcgagat	cccccggaac	caaaggaagt	
	10380						
	aggttaaacc	cgctccgatc	aggccgagcc	acgccaggcc	gagaacattg	gttctctgtg	
	10440						
35	gcatcgggat	tggcggatca	aacactaaag	ctactggaac	gagcagaagt	cctccggccg	
	10500						
	ccagttgcca	ggcggtaaag	gtgagcagag	gcacggggagg	ttgccacttg	cgggtcagca	
	10560						
	cggttccgaa	cgccatggaa	accgcccccg	ccaggccccg	tgcgacgccg	acaggatcta	
40	10620						
	gcgctgcgtt	tggtgtcaac	accaacagcg	ccacgccccg	agttccgcaa	atagccccca	
	10680						
	ggaccgccat	caatcgatc	gggctaccta	gcagagcggc	agagatgaac	acgaccatca	
	10740						
45	gcggctgcac	agcgcctacc	gtcgccgcga	ccccgccccg	caggcggtag	accgaaataa	
	10800						
	acaacaagct	ccagaatagc	gaaatattaa	gtgcgcgcgag	gatgaagatg	cgcattccacc	
	10860						
	agattccccg	tggaatctgt	cggacgatca	tcacgagcaa	taaaccgccc	ggcaacgccc	
50	10920						
	gcagcagcat	accggcgacc	cctcggcctc	gtgtttcggg	ctccacgaaa	acgccggaca	
	10980						
	gatgcgcctt	gtgagcgctc	ttggggccgt	cctcctgttt	gaagaccgac	agcccaatga	
	11040						
55	tctcgcgcgc	gatgtaggcg	ccgaatgcca	cggcatctcg	caaccgttca	gcgaacgcct	
	11100						
	ccatgggctt	tttctcctcg	tgctcgtaaa	cggacccgaa	catctctgga	gctttcttca	
	11160						
	gggccgacaa	tcggatctcg	cggaaatcct	gcacgtcggc	cgctccaagc	cgtcgaatct	
60	11220						

34

gagccttaat cacaattgtc aattttaate ctctgtttat cggcagttcg tagagcgcgc
 11280
 cgtgcgtccc gagcgatact gagcgaagca agtgcgtcga gcagtgcctcg cttgttcctg
 11340
 5 aaatgccagt aaagcgctgg ctgctgaacc cccagccgga actgacccca caaggcccta
 11400
 gcgtttgcaa tgcaccaggt catcattgac ccaggcgtgt tccaccaggc cgctgcctcg
 11460
 10 caactcttcg caggcttcgc cgacctgctc gcgccacttc ttcacgcggg tggaaatccga
 11520
 tccgcacatg aggcggaagg tttccagctt gagcgggtac ggctcccggg gcgagctgaa
 11580
 atagtgaac atccgtcggg ccgtcggcga cagcttgagg tacttctccc atatgaattt
 11640
 15 cgtgtagtgg tgcacagcaa acagcacgac gatttcctcg tcgatcagga cctggcaacg
 11700
 ggacgttttc ttgccacggg ccaggacgag gaagcggtgc agcagcgaca ccgattccag
 11760
 20 gtgcccacac cggtcggacg tgaagcccat cgccgtcgcc tgtaggcgag acaggcattc
 11820
 ctccgcttc gtgtaatacc ggccattgat cgaccagccc aggtcctggc aaagctcgta
 11880
 gaacgtgaag gtgatcggct cgccgatagg ggtgcgcttc gcgtactcca acacctgctg
 11940
 25 ccacaccagt tcgtcatcgt cggcccgcag ctcgacgccc gtgtaggtga tcttcacgtc
 12000
 cttgttgacg tggaaaatga ccttgttttg cagcgccctg cgcgggattt tcttgttgag
 12060
 30 cgtggtgaac agggcagagc gggccgtgtc gtttggcatc gctcgcatcg tgtccggcca
 12120
 cggcgcaata tcgaacaagg aaagctgcat ttccttgatc tgctgcttcg tgtgtttcag
 12180
 caacgcggcc tgcttggcct cgctgacctg ttttgccagg tctcgcggg cgggtttttc
 12240
 35 cttcttggc gtcatagtgc ctccgctgtc gatggtcacg gacttcgcca aacctgccc
 12300
 ctccgttcg agacgacgag aacgctccac ggccggccgat ggccggggga gggcaggggg
 12360
 40 agccagttgc acgctgtcgc gctcgatctt ggccgtagct tgctggacca tcgagccgac
 12420
 ggactggaag gtttcgaggg gcgcacgcat gacgggtgagg cttgcgatgg tttcggcatc
 12480
 ctccggcgga aaccccgct cgatcagttc ttgcctgtat gccttcggg caaacgtccg
 12540
 45 attcattcac cctccttgag ggattgcccc gactcacgac ggggcaatgt gcccttattc
 12600
 ctgatttgac ccgcttggtg ccttggtgtc cagataatcc acctatcgg caatgaagtc
 12660
 50 ggtcccgtag accgtctggc cgtccttctc gtacttggtg ttccgaatct tgccctgcac
 12720
 gaataccagc gaccttgc ccaaatactt gccgtgggac tcggcctgag agccaaaaca
 12780
 cttgatgcgg aagaagtcgg tgcgctcctg cttgtcgccg gcatcgttgc gccacatcta
 12840
 55 ggtactaaaa caattcatcc agtaaaatat aatattttat tttctcccaa tcaggcttga
 12900
 tccccagtaa gtcaaaaaat agctcgacat actgttcttc cccgatatcc tccctgatcg
 12960
 60 accggacgca gaaggcaatg tcataaccact tgtccgcctt gccgcttctc ccaagatcaa
 13020

taaagccact tacttttgcca tcttttcacaa agatgttgct gtctcccagg tcgccgtggg
 13080
 aaaagacaag ttccctcttcg ggctttttccg tcttttaaaaa atcatacagc tcgcgcggat
 13140
 5 cttttaaatgg agtgtcttct tcccagtttt cgcaatccac atcggccaga tcgttattca
 13200
 gtaagtaatc caattcggct aagcggtgt ctaagctatt cgtataggga caatccgata
 13260
 10 tgtcgatgga gtgaaagagc ctgatgcact ccgcatacag ctcgataatc ttttcagggc
 13320
 tttgttcac ttcatactct tccgagcaaa ggacgccatc ggcctcactc atgagcagat
 13380
 tgctccagcc atcatgccgt tcaaagtgc ggacctttgg aacaggcagc tttccttcca
 13440
 15 gccatagcat catgtccttt tcccgttcca catcatagggt ggtcccttta taccggctgt
 13500
 ccgtcatttt taaatatagg ttttcatttt cttccaccag cttatatacc ttagcaggag
 13560
 20 acattccttc cgtatctttt acgcagcgggt atttttcgat cagttttttc aattccgggtg
 13620
 atattctcat tttagccatt tattatttcc ttccctcttt ctacagtatt taaagatacc
 13680
 ccaagaagct aattataaca agacgaactc caattcactg ttccttgcat tctaaaacct
 13740
 25 taaataccag aaaacagctt tttcaaagtt gttttcaaag ttggcgtata acatagtatc
 13800
 gacggagccg attttgaaac cacaattatg ggtgatgctg ccaacttact gatttagtgt
 13860
 30 atgatgggtg ttttgagggt ctcagtggtc ttctgtgtct atcagctgtc cctcctgttc
 13920
 agctactgac ggggtggtgc gtaacggcaa aagcaccgcc ggacatcagc gctatctctg
 13980
 ctctcactgc cgtaaaacat ggcaactgca gttoacttac accgcttctc aaccgggtac
 14040
 35 gcaccagaaa atcattgata tggccatgaa tggcgttgga tgccgggcaa cagcccgc
 14100
 tatgggcgtt ggctcaaca cgattttacg tcaacttaaaa aactcaggcc gcagtcggta
 14160
 40 acctcgcgca tacagccggg cagtgcgtc atcgtctgcg cggaatgga cgaacagtgg
 14220
 ggctatgtcg gggctaaatc gcgccagcgc tggctgtttt acgcgtatga cagtctccgg
 14280
 aagacggttg ttgcgcacgt attcgggtgaa cgcactatgg cgacgctggg gcgtcttatg
 14340
 45 agcctgtgtg caccctttga cgtggtgata tggatgacgg atggctggcc gctgtatgaa
 14400
 tccgcctga agggaaagct gcacgtaatc agcaagcgat atacgcagcg aattgagcgg
 14460
 50 cataacctga atctgaggca gcacctggca cggctgggac ggaagtcgct gtcgttctca
 14520
 aaatcggttg agctgcatga caaagtcac gggcattatc tgaacataaa acactatcaa
 14580
 taagttggag tcattacca attatgatag aatttacaag ctataagggt attgtcctgg
 14640
 55 gtttcaagca ttagtccatg caagttttta tgctttgccc attctataga tatattgata
 14700
 agcgcgtgc ctatgccttg cccctgaaa tcttacata cggcgatatc ttctatataa
 14760
 60 aagatatatt atcttatcag tattgtcaat atattcaagg caatctgcct cctcatcctc
 14820

36

ttcatcctct tcgtcttggg agcttttttaa atatggcgct tcatagagta attctgtaaa
 14880
 ggtccaattc tcgttttcat acctcggtat aatcttacct atcacctcaa atggttcgct
 14940
 5 gggtttatcg ccccccgaa caccgagcacg gcacccgcga ccaactatgcc aagaatgccc
 15000
 aaggtaaaaa ttgccggccc cgccatgaag tccgtgaatg ccccgacggc cgaagtgaag
 15060
 10 ggcaggccgc caccaggcc gccgccctca ctgccggca cctggtcgct gaatgtcgat
 15120
 gccagcacct gcggcacgtc aatgcttccg ggcgtcgcgc tcgggctgat cgcccatccc
 15180
 gttactgccc cgatcccggc aatggcaagg actgccagcg ctgccatttt tggggtgagg
 15240
 15 ccgttcgchg ccgaggggch cagcccctgg ggggatggga ggcccgcgtt agcg
 15294

<210> 16
 20 <211> 2019
 <212> DNA
 <213> Arabidopsis thaliana

<220>
 25 <221> CDS
 <222> (1)..(2010)

<400> 16
 30 atg gcg gcg gca aca aca aca aca aca aca tct tct tcg atc tcc ttc 48
 Met Ala Ala Ala Thr Thr Thr Thr Thr Thr Ser Ser Ser Ile Ser Phe
 1 5 10 15

35 tcc acc aaa cca tct cct tcc tcc tcc aaa tca cca tta cca atc tcc 96
 Ser Thr Lys Pro Ser Pro Ser Ser Ser Lys Ser Pro Leu Pro Ile Ser
 20 25 30

40 aga ttc tcc ctc cca ttc tcc cta aac ccc aac aaa tca tcc tcc tcc 144
 Arg Phe Ser Leu Pro Phe Ser Leu Asn Pro Asn Lys Ser Ser Ser Ser
 35 40 45

45 tcc cgc cgc cgc ggt atc aaa tcc agc tct ccc tcc tcc atc tcc gcc 192
 Ser Arg Arg Arg Gly Ile Lys Ser Ser Ser Pro Ser Ser Ile Ser Ala
 50 55 60

50 gtg ctc aac aca acc acc aat gtc aca acc act ccc tct cca acc aaa 240
 Val Leu Asn Thr Thr Thr Asn Val Thr Thr Thr Pro Ser Pro Thr Lys
 65 70 75 80

55 cct acc aaa ccc gaa aca ttc atc tcc cga ttc gct cca gat caa ccc 288
 Pro Thr Lys Pro Glu Thr Phe Ile Ser Arg Phe Ala Pro Asp Gln Pro
 85 90 95

60 cgc aaa ggc gct gat att ctc gtc gag gct tta gaa cgt caa ggc gta 336
 Arg Lys Gly Ala Asp Ile Leu Val Glu Ala Leu Glu Arg Gln Gly Val
 100 105 110

60 gaa acc gta ttc gct tac cct gga ggt gca tca atg gag att cac caa 384
 Glu Thr Val Phe Ala Tyr Pro Gly Gly Ala Ser Met Glu Ile His Gln
 115 120 125

gcc tta acc cgc tct tcc tca atc cgt aac gtc ctt cct cgt cac gaa 432

37

	Ala	Leu	Thr	Arg	Ser	Ser	Ser	Ile	Arg	Asn	Val	Leu	Pro	Arg	His	Glu	
	130						135					140					
5	caa	gga	ggt	gta	ttc	gca	gca	gaa	gga	tac	gct	cga	tcc	tca	ggt	aaa	480
	Gln	Gly	Gly	Val	Phe	Ala	Ala	Glu	Gly	Tyr	Ala	Arg	Ser	Ser	Gly	Lys	
	145					150					155					160	
10	cca	ggt	atc	tgt	ata	gcc	act	tca	ggt	ccc	gga	gct	aca	aat	ctc	gtt	528
	Pro	Gly	Ile	Cys	Ile	Ala	Thr	Ser	Gly	Pro	Gly	Ala	Thr	Asn	Leu	Val	
					165					170					175		
15	agc	gga	tta	gcc	gat	gcg	ttg	tta	gat	agt	gtt	cct	ctt	gta	gca	atc	576
	Ser	Gly	Leu	Ala	Asp	Ala	Leu	Leu	Asp	Ser	Val	Pro	Leu	Val	Ala	Ile	
				180					185					190			
20	aca	gga	caa	gtc	cct	cgt	cgt	atg	att	ggt	aca	gat	gcg	ttt	caa	gag	624
	Thr	Gly	Gln	Val	Pro	Arg	Arg	Met	Ile	Gly	Thr	Asp	Ala	Phe	Gln	Glu	
			195					200					205				
25	act	ccg	att	gtt	gag	gta	acg	cgt	tcg	att	acg	aag	cat	aac	tat	ctt	672
	Thr	Pro	Ile	Val	Glu	Val	Thr	Arg	Ser	Ile	Thr	Lys	His	Asn	Tyr	Leu	
		210					215					220					
30	gtg	atg	gat	gtt	gaa	gat	att	cct	agg	att	att	gag	gag	gct	ttc	ttt	720
	Val	Met	Asp	Val	Glu	Asp	Ile	Pro	Arg	Ile	Ile	Glu	Glu	Ala	Phe	Phe	
		225				230					235					240	
35	tta	gct	act	tct	ggt	aga	cct	gga	cct	gtt	ttg	gtt	gat	gtt	cct	aaa	768
	Leu	Ala	Thr	Ser	Gly	Arg	Pro	Gly	Pro	Val	Leu	Val	Asp	Val	Pro	Lys	
				245						250					255		
40	gat	att	caa	caa	cag	ctt	gcg	att	cct	aat	tgg	gaa	cag	gct	atg	aga	816
	Asp	Ile	Gln	Gln	Gln	Leu	Ala	Ile	Pro	Asn	Trp	Glu	Gln	Ala	Met	Arg	
				260					265					270			
45	tta	cct	ggt	tat	atg	tct	agg	atg	cct	aaa	cct	ccg	gaa	gat	tct	cat	864
	Leu	Pro	Gly	Tyr	Met	Ser	Arg	Met	Pro	Lys	Pro	Pro	Glu	Asp	Ser	His	
			275					280					285				
50	ttg	gag	cag	att	gtt	agg	ttg	att	tct	gag	tct	aag	aag	cct	gtg	ttg	912
	Leu	Glu	Gln	Ile	Val	Arg	Leu	Ile	Ser	Glu	Ser	Lys	Lys	Pro	Val	Leu	
		290					295					300					
55	tat	gtt	ggt	ggt	ggt	tgt	ttg	aac	tct	agc	gat	gaa	ttg	ggt	agg	ttt	960
	Tyr	Val	Gly	Gly	Gly	Cys	Leu	Asn	Ser	Ser	Asp	Glu	Leu	Gly	Arg	Phe	
		305				310					315					320	
60	gtt	gag	ctt	acg	gga	atc	cct	gtt	gcg	agt	acg	ttg	atg	ggg	ctg	gga	1008
	Val	Glu	Leu	Thr	Gly	Ile	Pro	Val	Ala	Ser	Thr	Leu	Met	Gly	Leu	Gly	
					325					330					335		
65	tct	tat	cct	tgt	gat	gat	gag	ttg	tcg	tta	cat	atg	ctt	gga	atg	cat	1056
	Ser	Tyr	Pro	Cys	Asp	Asp	Glu	Leu	Ser	Leu	His	Met	Leu	Gly	Met	His	
				340					345					350			
70	ggg	act	gtg	tat	gca	aat	tac	gct	gtg	gag	cat	agt	gat	ttg	ttg	ttg	1104
	Gly	Thr	Val	Tyr	Ala	Asn	Tyr	Ala	Val	Glu	His	Ser	Asp	Leu	Leu	Leu	
			355					360					365				
75	gcg	ttt	ggg	gta	agg	ttt	gat	gat	cgt	gtc	acg	ggt	aaa	ctt	gag	gct	1152
	Ala	Phe	Gly	Val	Arg	Phe	Asp	Asp	Arg	Val	Thr	Gly	Lys	Leu	Glu	Ala	

38

	370	375	380	
5	ttt gct agt agg gct aag att gtt cat att gat att gac tcg gct gag Phe Ala Ser Arg Ala Lys Ile Val His Ile Asp Ile Asp Ser Ala Glu 385 390 395 400	1200		
10	att ggg aag aat aag act cct cat gtg tct gtg tgt ggt gat gtt aag Ile Gly Lys Asn Lys Thr Pro His Val Ser Val Cys Gly Asp Val Lys 405 410 415	1248		
15	ctg gct ttg caa ggg atg aat aag gtt ctt gag aac cga gcg gag gag Leu Ala Leu Gln Gly Met Asn Lys Val Leu Glu Asn Arg Ala Glu Glu 420 425 430	1296		
20	ctt aaa ctt gat ttt gga gtt tgg agg aat gag ttg aac gta cag aaa Leu Lys Leu Asp Phe Gly Val Trp Arg Asn Glu Leu Asn Val Gln Lys 435 440 445	1344		
25	cag aag ttt ccg ttg agc ttt aag acg ttt ggg gaa gct att cct cca Gln Lys Phe Pro Leu Ser Phe Lys Thr Phe Gly Glu Ala Ile Pro Pro 450 455 460	1392		
30	cag tat gcg att aag gtc ctt gat gag ttg act gat gga aaa gcc ata Gln Tyr Ala Ile Lys Val Leu Asp Glu Leu Thr Asp Gly Lys Ala Ile 465 470 475 480	1440		
35	ata agt act ggt gtc ggg caa cat caa atg tgg gcg gcg cag ttc tac Ile Ser Thr Gly Val Gly Gln His Gln Met Trp Ala Ala Gln Phe Tyr 485 490 495	1488		
40	aat tac aag aaa cca agg cag tgg cta tca tca gga ggc ctt gga gct Asn Tyr Lys Lys Pro Arg Gln Trp Leu Ser Ser Gly Gly Leu Gly Ala 500 505 510	1536		
45	atg gga ttt gga ctt cct gct gcg att gga gcg tct gtt gct aac cct Met Gly Phe Gly Leu Pro Ala Ala Ile Gly Ala Ser Val Ala Asn Pro 515 520 525	1584		
50	gat gcg ata gtt gtg gat att gac gga gat gga agt ttt ata atg aat Asp Ala Ile Val Val Asp Ile Asp Gly Asp Gly Ser Phe Ile Met Asn 530 535 540	1632		
55	gtg caa gag cta gcc act att cgt gta gag aat ctt cca gtg aag gta Val Gln Glu Leu Ala Thr Ile Arg Val Glu Asn Leu Pro Val Lys Val 545 550 555 560	1680		
60	ctt tta tta aac aac cag cat ctt ggc atg gtt atg caa tgg gaa gat Leu Leu Leu Asn Asn Gln His Leu Gly Met Val Met Gln Trp Glu Asp 565 570 575	1728		
65	cgg ttc tac aaa gct aac cga gca cac aca ttt ctc gga gat ccg gct Arg Phe Tyr Lys Ala Asn Arg Ala His Thr Phe Leu Gly Asp Pro Ala 580 585 590	1776		
70	cag gag gac gag ata ttc ccg aac atg ttg ctg ttt gca gca gct tgc Gln Glu Asp Glu Ile Phe Pro Asn Met Leu Leu Phe Ala Ala Ala Cys 595 600 605	1824		
75	ggg att cca gcg gcg agg gtg aca aag aaa gca gat ctc cga gaa gct Gly Ile Pro Ala Ala Arg Val Thr Lys Lys Ala Asp Leu Arg Glu Ala 610 615 620	1872		

att cag aca atg ctg gat aca cca gga cct tac ctg ttg gat gtg att 1920
 Ile Gln Thr Met Leu Asp Thr Pro Gly Pro Tyr Leu Leu Asp Val Ile
 625 630 635 640

5 tgt ccg cac caa gaa cat gtg ttg ccg atg atc ccg aat ggt ggc act 1968
 Cys Pro His Gln Glu His Val Leu Pro Met Ile Pro Asn Gly Gly Thr
 645 650 655

10 ttc aac gat gtc ata acg gaa gga gat ggc cgg att aaa tac tgagagctc 2019
 Phe Asn Asp Val Ile Thr Glu Gly Asp Gly Arg Ile Lys Tyr
 660 665 670

15 <210> 17
 <211> 670
 <212> PRT
 <213> Arabidopsis thaliana

20 <400> 17
 Met Ala Ala Ala Thr Thr Thr Thr Thr Thr Ser Ser Ser Ile Ser Phe
 1 5 10 15

25 Ser Thr Lys Pro Ser Pro Ser Ser Ser Lys Ser Pro Leu Pro Ile Ser
 20 25 30

Arg Phe Ser Leu Pro Phe Ser Leu Asn Pro Asn Lys Ser Ser Ser Ser
 35 40 45

30 Ser Arg Arg Arg Gly Ile Lys Ser Ser Ser Pro Ser Ser Ile Ser Ala
 50 55 60

Val Leu Asn Thr Thr Thr Asn Val Thr Thr Thr Pro Ser Pro Thr Lys
 65 70 75 80

35 Pro Thr Lys Pro Glu Thr Phe Ile Ser Arg Phe Ala Pro Asp Gln Pro
 85 90 95

40 Arg Lys Gly Ala Asp Ile Leu Val Glu Ala Leu Glu Arg Gln Gly Val
 100 105 110

Glu Thr Val Phe Ala Tyr Pro Gly Gly Ala Ser Met Glu Ile His Gln
 115 120 125

45 Ala Leu Thr Arg Ser Ser Ser Ile Arg Asn Val Leu Pro Arg His Glu
 130 135 140

Gln Gly Gly Val Phe Ala Ala Glu Gly Tyr Ala Arg Ser Ser Gly Lys
 145 150 155 160

50 Pro Gly Ile Cys Ile Ala Thr Ser Gly Pro Gly Ala Thr Asn Leu Val
 165 170 175

55 Ser Gly Leu Ala Asp Ala Leu Leu Asp Ser Val Pro Leu Val Ala Ile
 180 185 190

Thr Gly Gln Val Pro Arg Arg Met Ile Gly Thr Asp Ala Phe Gln Glu
 195 200 205

60 Thr Pro Ile Val Glu Val Thr Arg Ser Ile Thr Lys His Asn Tyr Leu
 210 215 220

	Val	Met	Asp	Val	Glu	Asp	Ile	Pro	Arg	Ile	Ile	Glu	Glu	Ala	Phe	Phe	225	230	235	240
5	Leu	Ala	Thr	Ser	Gly	Arg	Pro	Gly	Pro	Val	Leu	Val	Asp	Val	Pro	Lys	245	250	255	
	Asp	Ile	Gln	Gln	Gln	Leu	Ala	Ile	Pro	Asn	Trp	Glu	Gln	Ala	Met	Arg	260	265	270	
10	Leu	Pro	Gly	Tyr	Met	Ser	Arg	Met	Pro	Lys	Pro	Pro	Glu	Asp	Ser	His	275	280	285	
15	Leu	Glu	Gln	Ile	Val	Arg	Leu	Ile	Ser	Glu	Ser	Lys	Lys	Pro	Val	Leu	290	295	300	
	Tyr	Val	Gly	Gly	Gly	Cys	Leu	Asn	Ser	Ser	Asp	Glu	Leu	Gly	Arg	Phe	305	310	315	320
20	Val	Glu	Leu	Thr	Gly	Ile	Pro	Val	Ala	Ser	Thr	Leu	Met	Gly	Leu	Gly	325	330	335	
	Ser	Tyr	Pro	Cys	Asp	Asp	Glu	Leu	Ser	Leu	His	Met	Leu	Gly	Met	His	340	345	350	
25	Gly	Thr	Val	Tyr	Ala	Asn	Tyr	Ala	Val	Glu	His	Ser	Asp	Leu	Leu	Leu	355	360	365	
30	Ala	Phe	Gly	Val	Arg	Phe	Asp	Asp	Arg	Val	Thr	Gly	Lys	Leu	Glu	Ala	370	375	380	
	Phe	Ala	Ser	Arg	Ala	Lys	Ile	Val	His	Ile	Asp	Ile	Asp	Ser	Ala	Glu	385	390	395	400
35	Ile	Gly	Lys	Asn	Lys	Thr	Pro	His	Val	Ser	Val	Cys	Gly	Asp	Val	Lys	405	410	415	
	Leu	Ala	Leu	Gln	Gly	Met	Asn	Lys	Val	Leu	Glu	Asn	Arg	Ala	Glu	Glu	420	425	430	
40	Leu	Lys	Leu	Asp	Phe	Gly	Val	Trp	Arg	Asn	Glu	Leu	Asn	Val	Gln	Lys	435	440	445	
45	Gln	Lys	Phe	Pro	Leu	Ser	Phe	Lys	Thr	Phe	Gly	Glu	Ala	Ile	Pro	Pro	450	455	460	
	Gln	Tyr	Ala	Ile	Lys	Val	Leu	Asp	Glu	Leu	Thr	Asp	Gly	Lys	Ala	Ile	465	470	475	480
50	Ile	Ser	Thr	Gly	Val	Gly	Gln	His	Gln	Met	Trp	Ala	Ala	Gln	Phe	Tyr	485	490	495	
	Asn	Tyr	Lys	Lys	Pro	Arg	Gln	Trp	Leu	Ser	Ser	Gly	Gly	Leu	Gly	Ala	500	505	510	
55	Met	Gly	Phe	Gly	Leu	Pro	Ala	Ala	Ile	Gly	Ala	Ser	Val	Ala	Asn	Pro	515	520	525	
60	Asp	Ala	Ile	Val	Val	Asp	Ile	Asp	Gly	Asp	Gly	Ser	Phe	Ile	Met	Asn	530	535	540	

41

Val Gln Glu Leu Ala Thr Ile Arg Val Glu Asn Leu Pro Val Lys Val
 545 550 555 560

5 Leu Leu Leu Asn Asn Gln His Leu Gly Met Val Met Gln Trp Glu Asp
 565 570 575

Arg Phe Tyr Lys Ala Asn Arg Ala His Thr Phe Leu Gly Asp Pro Ala
 580 585 590

10 Gln Glu Asp Glu Ile Phe Pro Asn Met Leu Leu Phe Ala Ala Ala Cys
 595 600 605

Gly Ile Pro Ala Ala Arg Val Thr Lys Lys Ala Asp Leu Arg Glu Ala
 610 615 620

15 Ile Gln Thr Met Leu Asp Thr Pro Gly Pro Tyr Leu Leu Asp Val Ile
 625 630 635 640

20 Cys Pro His Gln Glu His Val Leu Pro Met Ile Pro Asn Gly Gly Thr
 645 650 655

Phe Asn Asp Val Ile Thr Glu Gly Asp Gly Arg Ile Lys Tyr
 660 665 670

25

<210> 18
 <211> 259
 <212> DNA
 30 <213> Artificial Sequence

<220>
 <221> promoter
 <222> (1)..(259)

35

<220>
 <223> Description of Artificial Sequence:be2promoter
 fragment

40 <400> 18
 gatctcctaaa taattcgaaa tatctttgtt attatttttt tctattcaaa ttgcaattag 60
 acataagtca ttttaactga agctgcattg atgaaaaatt atactatgtc tttatgtata 120
 tatattaatg ttttaaatc ctttatagtg ataaagatgg ttcgaaacat gctacaaatt 180
 attatacgaa gttacttttt ttaatctact ttaacaattt tctaatttca ctattgaaca 240
 45 tagataccag cccgggccg 259

<210> 19
 <211> 400
 50 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:RNAi420be2be1

55

<220>

<400> 19
 60 gaattgttgt tctcatggac atcgttcaca gccatgcac aaataatact ttagatggac 60
 tgaacatgtt tgacggcacc gatagttgtt actttcactc tggagctcgt gggtatcatt 120
 ggatgtggga ttcccgcctc ttttaactatg gaaactggga ggtacttagg tatcttctct 180

caaatgcgag atggtggttg ccatttcaca tcaccagaag gaatacctgg agttccagaa 240
 acaaatttca atggtcgtcc aaattccttc aaagtgtgtg ctctgctgcg aacatgtgtg 300
 gcttattaca gagttgacga acgcatgtca gaaactgaag tttaccagac agacatttct 360
 agtgagctac taccaacagc caatatcgag gagagtgcg 400

5

<210> 20
 <211> 1105
 <212> DNA
 10 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:RNAifragment

15 <220>

<400> 20
 gaattgttgt tctcatggac atcgttcaca gccatgcac aaataatact ttagatggac 60
 tgaacatgtt tgacggcacc gatagtgtt actttcactc tggagctcgt gggtatcatt 120
 20 ggatgtggga ttccgcctc tttaactatg gaaactggga ggtacttagg tatcttctct 180
 caaatgcgag atggtggttg ccatttcaca tcaccagaag gaatacctgg agttccagaa 240
 acaaatttca atggtcgtcc aaattccttc aaagtgtgtg ctctgctgcg aacatgtgtg 300
 gcttattaca gagttgacga acgcatgtca gaaactgaag tttaccagac agacatttct 360
 agtgagctac taccaacagc caatatcgag gagagtgcg atcaagctga tctctaaata 420
 25 attcgaaata tctttgttat tatttttttc tattcaaatt gcaattagac ataagtcatt 480
 ttaactgaag ttgcattgat gaaaaattat actatgtttt atgtatatat attaatTTTT 540
 aaattccttt atagtataa agatagttcg aaacatgcta taaattatta tacgaattta 600
 cgttactttt tttaatctac tttaacaatt ttctaatttc actattgaac atagatacca 660
 gcccgggccg tcgacctga attcgccctt ggagagtgc gttcgcgtca ctctcctcga 720
 30 tattggctgt tggtagtagc tcactagaaa tgtctgtctg gtaaaacttca gtttctgaca 780
 tgcgttcgtc aactctgtaa taagccacac atgttcgcgc aggagacagc actttgaagg 840
 aatttgagcg accattgaaa tttgtttctg gaactccagg tattccttct ggtgatgtga 900
 aatggcaacc accatctcgc atttgagaga agatacctaa gtacctccca gtttccatag 960
 ttaaagaggc gggaatccca catccaatga taaccacgag ctccagagtg aaagtaacaa 1020
 35 ctatcggtgc cgtcaaacat gttcagtcga tctaaagtat tatttgatgc atggctgtga 1080
 acgatgtcca tgagaacaac aattc 1105

<210> 21
 40 <211> 180
 <212> DNA
 <213> Artificial Sequence

<220>
 45 <223> Description of Artificial Sequence:SBE RNAi 1

<220>
 <400> 21
 50 actagtggta cttaggtatc ttctctcaaa tgcgagatgg tggttgagtg agctactacc 60
 aacagccaat atcgaggaga gtgacgttcg cgtcactctc ctcgatattg gctgttggtg 120
 gtagctcact caaccacat ctgcgatttg agagaagata cctaagtacc ttttgggtacc 180

55 <210> 22
 <211> 420
 <212> DNA
 <213> Artificial Sequence

60 <220>
 <223> Description of Artificial Sequence:SBE RNAi 2

<400> 22
 actagttgga gctcgtgggt atcattggat gtgggattcc cgccctcttta actatggaaa 60
 ctgggaggta cttagggtatc ttctctcaaa tgcgagatgg tggttggctt attacagagt 120
 5 tgacgaacgc atgtcagaaa ctgaagttta ccagacagac atttctagtg agctactacc 180
 aacagccaat atcgaggaga gtgacgttcg cgctactctc ctcgatattg gctgttggtta 240
 gtagctcact agaaatgtct gtctggtaaa cttcagtttc tgacatgcgt tcgtcaactc 300
 tgtaataagc caaccaccat ctgcgatttg agagaagata cctaagtacc tccagtttc 360
 10 catagttaaa gaggcgggaa tcccacatcc aatgataacc acgagctcca ttttgggtacc 420

<210> 23
 <211> 37
 <212> DNA
 15 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: spacer

20 <220>

<400> 23
 atcaagctta tcgataccgt cgacctcgaa gcttgat 37

25

<210> 24
 <211> 837
 <212> DNA
 <213> Artificial Sequence

30

<220>
 <223> Description of Artificial Sequence: fragment of be2
 and bel in pHAS3 for RNAi

35 <220>

<400> 24
 gaattgttgt tctcatggac atcgttcaca gccatgcac aaataatact ttagatggac 60
 tgaacatgtt tgacggcacc gatagttgtt actttcactc tggagctcgt ggttatcatt 120
 40 ggatgtggga ttcccgcctc tttaactatg gaaactggga ggtacttagg tatcttctct 180
 caaatgcgag atggtggttg ccatttcaca tcaccagaag gaataacctg agttccagaa 240
 acaaatttca atggtcgtcc aaattccttc aaagtgctgt ctccctgcgcg aacatgtgtg 300
 gcttattaca gagttgacga acgcatgtca gaaactgaag tttaccagac agacatttct 360
 agtgagctac taccaacagc caatatcgag gagagtgaag atcaagctta tcgataccgt 420
 45 cgacctcgaa gcttgatcgt cactctcctc gatattggct gttggtagta gctcactaga 480
 aatgtctgtc tggtaaactt cagtttctga catgcgttcg tcaactctgt aataagccac 540
 acatgttcgc gcaggagaca gcactttgaa ggaatttgga cgaccattga aatttggttc 600
 tggaactcca ggtattcctt ctggtgatgt gaaatggcaa ccaccatctc gcatttgaga 660
 gaagatacct aagtacctcc cagtttccat agttaagag gcgggaatcc cacatccaat 720
 50 gataaccacg agctccagag tgaaagtaac aactatcggt gccgtcaaac atgttcagtc 780
 catctaaagt attatttgat gcatggctgt gaacgatgtc catgagaaca acaattc 837

This Page Blank (uspto)